

1309.43471X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: I. URATANI, et al

Serial No.: 10/769,784

Filing Date: February 3, 2004

For: DISK ARRAY APPARATUS, AND METHOD FOR CONTROLLING THE  
SAME

**LETTER CLAIMING RIGHT OF PRIORITY**

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

May 11, 2004

Sir:


Under the provisions of 35 USC 119 and 37 CFR 1.55, applicants hereby claim  
the right of priority based on:

**Japanese Application No. 2003-400549  
Filed: November 28, 2003**

A Certified copy of said application document is attached hereto.

Acknowledgment thereof is respectfully requested.

Respectfully submitted,

  
\_\_\_\_\_  
Carl I. Brundidge  
Registration No. 29,621  
ANTONELLI, TERRY, STOUT & KRAUS, LLP

CIB/jdc  
Enclosures  
703/312-6600

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日                    2 0 0 3 年 1 1 月 2 8 日  
Date of Application:

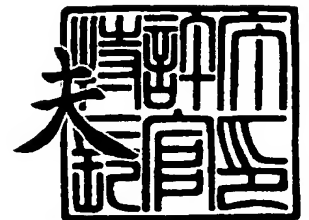
出 願 番 号                    特 願 2 0 0 3 - 4 0 0 5 4 9  
Application Number:  
[ST. 10/C] :                    [ J P 2 0 0 3 - 4 0 0 5 4 9 ]

出      願      人                    株 式 会 社 日 立 製 作 所  
Applicant(s):

2 0 0 4 年    1 月 2 7 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康



出証番号    出証特 2 0 0 4 - 3 0 0 2 9 6 3

【書類名】 特許願  
【整理番号】 340301169  
【あて先】 特許庁長官殿  
【国際特許分類】 G06F 03/06  
【発明者】  
    【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I  
                            D システム事業部内  
    【氏名】 裏谷 郁夫  
【発明者】  
    【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 R A I  
                            D システム事業部内  
    【氏名】 占部 喜一郎  
【特許出願人】  
    【識別番号】 000005108  
    【氏名又は名称】 株式会社日立製作所  
【代理人】  
    【識別番号】 100095371  
    【弁理士】  
    【氏名又は名称】 上村 輝之  
【選任した代理人】  
    【識別番号】 100089277  
    【弁理士】  
    【氏名又は名称】 宮川 長夫  
【選任した代理人】  
    【識別番号】 100104891  
    【弁理士】  
    【氏名又は名称】 中村 猛  
【手数料の表示】  
    【予納台帳番号】 043557  
    【納付金額】 21,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1  
    【包括委任状番号】 0110323

**【書類名】 特許請求の範囲****【請求項 1】**

ホストコンピュータ及び／又は前記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、前記他のディスクアレイ装置の有する記憶領域と該自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、前記ホストコンピュータから送られてくるデータを記憶するディスクアレイ装置において、

該自ディスクアレイ装置が有する記憶領域と前記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して前記ホストコンピュータからのアクセス対象とする論理ユニット形成部と、

前記ホストコンピュータ及び／又は前記他のディスクアレイ装置から転送されるデータを、前記各論理ユニットにライトするデータライト部と、

前記ホストコンピュータから自ディスクアレイ装置及び／又は前記他のディスクアレイ装置に格納されているデータのリード要求があった場合に、該データリード要求の正当性をチェックする第 1 のチェック部と、

前記第 1 のチェック部が前記データリード要求が正当であると判断した場合に、前記データリード要求に基づいて前記各論理ユニットに保存されているデータを前記ホストコンピュータへ転送するデータ転送部と、

前記ホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、該指示の正当性をチェックする第 2 のチェック部と、

前記第 2 のチェック部によるチェックの結果、前記ホストコンピュータからの前記指示が正当であった場合に、前記指示に応じて前記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成部と、

前記ペア状態を形成している 2 つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して前記ホストコンピュータからのライトアクセスがあった場合に、該論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするコピー部と、

を備え、

前記コピー部が、前記複数の論理ユニットについての情報、及び前記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、前記他のディスクアレイ装置が有する実際の記憶領域に対して前記ライトデータをコピーするようにしたディスクアレイ装置。

**【請求項 2】**

請求項 1 記載のディスクアレイ装置において、

前記データリード要求が、自ディスクアレイ装置に対する前記他のディスクアレイ装置に格納されているデータの取得を指示するためのコマンドと、このコマンドに基づいて自ディスクアレイ装置が前記他のディスクアレイ装置から取得したデータの前記ホストコンピュータへの転送を指示するコマンドと、を含むディスクアレイ装置。

**【請求項 3】**

請求項 1 又は請求項 2 記載のディスクアレイ装置において、

前記第 1 のチェック部が、前記データ取得を指示するためのコマンド、及び前記ホストコンピュータへのデータの転送を指示するコマンドの双方の正当性を夫々チェックするディスクアレイ装置。

**【請求項 4】**

請求項 3 記載のディスクアレイ装置において、

前記第 1 のチェック部が、前記各コマンドをチェックした結果、双方又は何れか一方の

コマンドに正当性が無いと判断した場合には、夫々エラーが生じた旨を前記ホストコンピュータへ報告するディスクアレイ装置。

【請求項5】

請求項1乃至請求項4のいずれか1項記載のディスクアレイ装置において、

前記第1のチェック部が前記データの取得を指示するためのコマンドに正当性があると判断した場合に、前記ホストコンピュータへ転送すべきデータを該コマンドに基づいて作成するデータ作成部、を更に備えるディスクアレイ装置。

【請求項6】

請求項1乃至請求項5のいずれか1項記載のディスクアレイ装置において、

前記データ転送部が、前記データ作成部により作成されたデータを、前記ホストコンピュータが前記データリード要求の送信元であることを確認してから転送するようにしたディスクアレイ装置。

【請求項7】

請求項1記載のディスクアレイ装置において、

前記データライト部が、前記他のディスクアレイ装置から転送されるデータを、前記論理ユニットに接続される仮想化されたデバイスである、自ディスクアレイ装置が有する記憶領域にマッピングすることにより自ディスクアレイ装置へ書き込むようにしたディスクアレイ装置。

【請求項8】

ホストコンピュータ及び／又は前記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、前記他のディスクアレイ装置の有する記憶領域と該自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、前記ホストコンピュータから送られてくるデータを記憶するディスクアレイ装置において、

該自ディスクアレイ装置が有する記憶領域と前記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して前記ホストコンピュータからのアクセス対象とする論理ユニット形成部と、

自ディスクアレイ装置が有する記憶領域を仮想化した仮想デバイスの識別情報を少なくとも含み、且つ、前記仮想デバイスと前記仮想デバイスにマッピングされた前記他のディスクアレイ装置が有する記憶領域との対応関係を示すマッピングテーブルを保持するマッピングテーブル保持部と、

前記ホストコンピュータから少なくとも前記仮想デバイスの識別情報を含むデータリード要求が送信された場合に、前記識別情報に基づいて対象となる仮想デバイスを前記マッピングテーブルから検索する検索部と、

前記検索部によって検索された仮想デバイスからデータをリードして前記ホストコンピュータへ転送するデータ転送部と、

前記ホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、該指示の正当性をチェックするチェック部と、

前記チェック部によるチェックの結果、前記ホストコンピュータからの前記指示が正当であった場合に、前記指示に応じて前記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成部と、

前記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して前記ホストコンピュータからのライトアクセスがあった場合に、該論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするコピー部と、

を備え、

前記コピー部が、前記複数の論理ユニットについての情報、及び前記他のディスクアレ

イ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、前記他のディスクアレイ装置が有する実際の記憶領域に対して前記ライトデータをコピーするようにしたディスクアレイ装置。

【請求項 9】

請求項 8 記載のディスクアレイ装置において、

前記仮想デバイスには、自ディスクアレイ装置が有する別の記憶領域のデータもコピーされるようになっており、

前記ホストコンピュータからのデータリード要求には、前記別の記憶領域の識別情報も含まれるディスクアレイ装置。

【請求項 10】

請求項 8 記載のディスクアレイ装置において、

前記データ転送部によって自ディスクアレイ装置から前記ホストコンピュータに転送されたデータが、前記ホストコンピュータから自ディスクアレイ装置には直接に接続されることが無い別のホストコンピュータへ更に転送されるディスクアレイ装置。

【請求項 11】

ホストコンピュータ及び／又は前記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、前記他のディスクアレイ装置の有する記憶領域と該自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、前記ホストコンピュータから送られてくるデータを記憶するディスクアレイ装置において、

該自ディスクアレイ装置が有する記憶領域と前記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して前記ホストコンピュータからのアクセス対象とする論理ユニット形成部と、

前記ホストコンピュータ及び／又は前記他のディスクアレイ装置から転送されるデータを、前記各論理ユニットにライトするデータライト部と、

前記ホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、該指示の正当性をチェックするチェック部と、

前記チェック部によるチェックの結果、前記ホストコンピュータからの前記指示が正当であった場合に、前記指示に応じて前記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成部と、

前記ペア状態を形成している 2 つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して前記ホストコンピュータからのライトアクセスがあった場合に、該論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするコピー部と、

を備え、

前記コピー部が、前記複数の論理ユニットについての情報、及び前記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、前記他のディスクアレイ装置が有する実際の記憶領域に対して前記ライトデータをコピーするようにしたディスクアレイ装置。

【請求項 12】

請求項 11 記載のディスクアレイ装置において、

前記チェック部が、前記ペア作成の指示の正当性をチェックした結果、正当性が無いと判断した場合には、エラーが生じた旨を前記ホストコンピュータへ報告するようにしたディスクアレイ装置。

【請求項 13】

請求項 11 記載のディスクアレイ装置において、

前記ペア状態形成部が、前記ペア状態形成の処理において問題が生じたと判断した場合

には、エラーが生じた旨を前記ホストコンピュータへ報告するようにしたディスクアレイ装置。

【請求項 14】

第1のホストコンピュータ及び／又は第2のホストコンピュータに接続された他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、前記他のディスクアレイ装置が有する記憶領域と該自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、前記第1のホストコンピュータから送られてくるデータを記憶するディスクアレイ装置において、

該自ディスクアレイ装置が有する記憶領域と前記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して前記ホストコンピュータからのアクセス対象とする論理ユニット形成部と、

前記ホストコンピュータ及び／又は前記他のディスクアレイ装置から転送されるデータを、前記各論理ユニットにライトするデータライト部と、

前記第1のホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、該指示の正当性をチェックするチェック部と、

前記チェック部によるチェックの結果、前記第1のホストコンピュータからの前記指示が正当であった場合に、前記指示に応じて前記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成部と、

前記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して前記第1のホストコンピュータからのライトアクセスがあった場合に、該論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするコピー部と、

前記第2のホストコンピュータから自ディスクアレイ装置及び／又は前記他のディスクアレイ装置に格納されているデータのリード要求があった場合に、前記データリード要求に基づく処理が開始されてから終了するまでの間、前記第2のホストコンピュータによる前記他のディスクアレイ装置へのアクセスを禁止する禁止部と、

を備え、

前記コピー部が、前記複数の論理ユニットについての情報、及び前記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、前記他のディスクアレイ装置が有する実際の記憶領域に対して前記ライトデータをコピーするようにしたディスクアレイ装置。

【請求項 15】

請求項 14 記載のディスクアレイ装置において、

前記アクセス禁止が、前記第2のホストコンピュータによる前記他のディスクアレイ装置へのアクセスから前記他のディスクアレイ装置のデータを保護するための前記禁止部による前記他のディスクアレイ装置へのリザーブコマンドの発行であるディスクアレイ装置。

【請求項 16】

請求項 14 記載のディスクアレイ装置において、

前記仮想デバイスには、自ディスクアレイ装置が有する別の記憶領域のデータもコピーされるようになっているディスクアレイ装置。

【請求項 17】

ホストコンピュータ及び／又は前記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、前記他のディスクアレイ装置の有する記憶領域と該自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、前記ホストコンピュータから送られてくるデータを記憶するディスクアレイ

装置の制御方法において、

該自ディスクアレ装置が有する記憶領域と前記他のディスクアレ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して前記ホストコンピュータからのアクセス対象とする論理ユニット形成のステップと、

前記ホストコンピュータ及び／又は前記他のディスクアレ装置から転送されるデータを、前記各論理ユニットにライトするデータライトのステップと、

前記ホストコンピュータから自ディスクアレ装置及び／又は前記他のディスクアレ装置に格納されているデータのリード要求があった場合に、該データリード要求の正当性をチェックする第1のチェックのステップと、

前記第1のチェックのステップにおいて前記データリード要求が正当であると判断した場合に、前記データリード要求に基づいて前記各論理ユニットに保存されているデータを前記ホストコンピュータへ転送するデータ転送のステップと、

前記ホストコンピュータから、自ディスクアレ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、該指示の正当性をチェックする第2のチェックのステップと、

前記第2のチェックのステップにおけるチェックの結果、前記ホストコンピュータからの前記指示が正当であった場合に、前記指示に応じて前記複数の論理ユニットのうち、自ディスクアレ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成のステップと、

前記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレ装置が有する記憶領域から形成された論理ユニットに対して前記ホストコンピュータからのライトアクセスがあった場合に、該論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするライトデータコピーのステップと、

を備え、

前記ライトデータコピーのステップが、前記複数の論理ユニットについての情報、及び前記他のディスクアレ装置が有する記憶領域と自ディスクアレ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、前記他のディスクアレ装置が有する実際の記憶領域に対して前記ライトデータをコピーするようにしたディスクアレ装置の制御方法。

#### 【請求項18】

ホストコンピュータ及び／又は前記ホストコンピュータとは直接接続できない他のディスクアレ装置と通信可能に接続されるディスクアレ装置であって、前記他のディスクアレ装置の有する記憶領域と該自ディスクアレ装置が有する記憶領域とを併せた記憶領域に対して、前記ホストコンピュータから送られてくるデータを記憶するディスクアレ装置の制御方法において、

該自ディスクアレ装置が有する記憶領域と前記他のディスクアレ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して前記ホストコンピュータからのアクセス対象とする論理ユニット形成のステップと、

自ディスクアレ装置が有する記憶領域を仮想化した仮想デバイスの識別情報を少なくとも含み、且つ、前記仮想デバイスと前記仮想デバイスにマッピングされた前記他のディスクアレ装置が有する記憶領域との対応関係を示すマッピングテーブルを保持するマッピングテーブル保持のステップと、

前記ホストコンピュータから少なくとも前記仮想デバイスの識別情報を含むデータリード要求が送信された場合に、前記識別情報に基づいて対象となる仮想デバイスを前記マッピングテーブルから検索する検索のステップと、

前記検索のステップにおいて検索された仮想デバイスからデータをリードして前記ホストコンピュータへ転送するデータ転送のステップと、

前記ホストコンピュータから、自ディスクアレ装置が有する記憶領域から形成された



論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、該指示の正当性をチェックするチェックのステップと、

前記チェックのステップにおけるチェックの結果、前記ホストコンピュータからの前記指示が正当であった場合に、前記指示に応じて前記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成のステップと、

前記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して前記ホストコンピュータからのライトアクセスがあった場合に、該論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするライトデータコピーのステップと、

を備え、

前記ライトデータコピーのステップが、前記複数の論理ユニットについての情報、及び前記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、前記他のディスクアレイ装置が有する実際の記憶領域に対して前記ライトデータをコピーするようにしたディスクアレイ装置の制御方法。

#### 【請求項19】

ホストコンピュータ及び／又は前記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、前記他のディスクアレイ装置の有する記憶領域と該自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、前記ホストコンピュータから送られてくるデータを記憶するディスクアレイ装置の制御方法において、

該自ディスクアレイ装置が有する記憶領域と前記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して前記ホストコンピュータからのアクセス対象とする論理ユニット形成のステップと、

前記ホストコンピュータ及び／又は前記他のディスクアレイ装置から転送されるデータを、前記各論理ユニットにライトするデータライトのステップと、

前記ホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、該指示の正当性をチェックするチェックのステップと、

前記チェックのステップにおけるチェックの結果、前記ホストコンピュータからの前記指示が正当であった場合に、前記指示に応じて前記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成のステップと、

前記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して前記ホストコンピュータからのライトアクセスがあった場合に、該論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするライトデータコピーのステップと、

を備え、

前記ライトデータコピーのステップが、前記複数の論理ユニットについての情報、及び前記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、前記他のディスクアレイ装置が有する実際の記憶領域に対して前記ライトデータをコピーするようにしたディスクアレイ装置の制御方法。

#### 【請求項20】

第1のホストコンピュータ及び／又は第2のホストコンピュータに接続された他のディス

クアレイ装置と通信可能に接続されるディスクアレイ装置であって、前記他のディスクアレイ装置が有する記憶領域と該自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、前記第1のホストコンピュータから送られてくるデータを記憶するディスクアレイ装置の制御方法において、

該自ディスクアレイ装置が有する記憶領域と前記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して前記ホストコンピュータからのアクセス対象とする論理ユニット形成のステップと、

前記ホストコンピュータ及び／又は前記他のディスクアレイ装置から転送されるデータを、前記各論理ユニットにライトするデータライトのステップと、

前記第1のホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、該指示の正当性をチェックするチェックのステップと、

前記チェックのステップにおけるチェックの結果、前記第1のホストコンピュータからの前記指示が正当であった場合に、前記指示に応じて前記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと前記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成のステップと、

前記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して前記第1のホストコンピュータからのライトアクセスがあった場合に、該論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするライトデータコピーのステップと、

前記第2のホストコンピュータから自ディスクアレイ装置及び／又は前記他のディスクアレイ装置に格納されているデータのリード要求があった場合に、前記データリード要求に基づく処理が開始されてから終了するまでの間、前記第2のホストコンピュータによる前記他のディスクアレイ装置へのアクセスを禁止するアクセス禁止のステップと、

を備え、

前記ライトデータコピーのステップが、前記複数の論理ユニットについての情報、及び前記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、前記他のディスクアレイ装置が有する実際の記憶領域に対して前記ライトデータをコピーするようにしたディスクアレイ装置の制御方法。

## 【書類名】明細書

【発明の名称】 ディスクアレイ装置、及びディスクアレイ装置の制御方法

## 【技術分野】

## 【0001】

本発明は、ホストコンピュータ及び／又は上記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、上記他のディスクアレイ装置の有する記憶領域とこの自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、上記ホストコンピュータから送られてくるデータを記憶するディスクアレイ装置、及びその制御方法に関する。なお、以下では、ディスクアレイ装置を記憶制御装置と表記して説明する。

## 【背景技術】

## 【0002】

例えば、データセンタ等のような大規模なデータを取り扱うデータベースシステムでは、ホストコンピュータ（以下、「ホスト装置」と表記する）とは別に構成された記憶システムを用いてデータを管理する。この記憶システムは、例えば、ディスクアレイ装置等から構成される。ディスクアレイ装置は、多数の記憶デバイスをアレイ状に配設して構成されるもので、例えば、RAID (Redundant Array of Independent Inexpensive Disks) に基づいて構築されている。記憶デバイス群が提供する物理的な記憶領域上には少なくとも1つ以上の論理ボリューム（論理ユニット）が形成され、この論理ボリュームがホスト装置（より詳しくは、ホスト装置上で稼働するデータベースプログラム（以下、「アプリケーションプログラム」ともいう））に提供される。ホスト装置は、所定のコマンドを送信することにより、論理ボリュームに対してデータの書き込み、読み出しを行うことができる。

## 【0003】

情報化社会の進展等につれて、データベースで管理すべきデータは、日々増大する。このため、より高性能、より大容量の記憶制御装置が求められており、この市場要求に応えるべく、新型の記憶制御装置が開発されている。新型の記憶制御装置を記憶システムを導入する方法としては、2つ考えられる。その一つは、旧型の記憶制御装置と新型の記憶制御装置とを完全に入れ替え、全て新型の記憶制御装置から記憶システムを構成する方法である（特許文献1）。他の一つは、旧型の記憶制御装置からなる記憶システムに新型の記憶制御装置を新たに追加し、新旧の記憶制御装置を併存させる方法である。

## 【0004】

なお、物理デバイスの記憶領域をセクタ単位で管理し、論理デバイスをセクタ単位で動的に構成する技術も知られている（特許文献2）。

## 【0005】

更に、容量の異なる複数の記憶デバイスから論理デバイスを構築する際に、最も容量の少ない記憶デバイスに合わせてエリアを形成し、残りの容量も最も小さな容量に合わせてエリアを形成するようにした技術も知られている（特許文献3）。

## 【0006】

【特許文献1】 特表平10-508967号公報

【特許文献2】 特開2001-337850号公報

【特許文献3】 特開平9-288547号公報

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0007】

ところで、ホスト装置上のアプリケーションプログラムが、ホスト装置から見て外部にある記憶制御装置からのインクアイリ（照会）情報を取得しようとするに際して、ホスト装置と外部の記憶制御装置との間の接続が不良なため、ホスト装置と外部の記憶制御装置との間のデータの通信性能が悪い場合がある。また、上記インクアイリ情報の取得に際して、ホスト装置と外部の記憶制御装置との間を接続することができないため、アプリケー

ションプログラムが外部の記憶制御装置内のインクアイリ情報をサポートできない場合もある。

【0008】

また、ホスト装置、及び該ホスト装置に接続される内部の記憶制御装置から見て外部のホスト装置（のアプリケーションプログラム）が、同じく外部の記憶制御装置からのインクアイリ情報を取得しようとする際に、外部の記憶制御装置の構成が、同じく外部のホスト装置からの（リモートコピーの）コマンドに対応不能な場合もある。

【0009】

更に、ホスト装置（のアプリケーションプログラム）から内部の記憶制御装置に発行されたコピー指示のコマンドに従って、内部の記憶制御装置が、リモートコピー非対応の外部記憶制御装置の記憶デバイスから自装置内の仮想デバイスへのリモートコピーを実行する場合がある。また、このリモートコピーと共に、内部の記憶制御装置が、自装置内の記憶デバイスから仮想デバイスへのコピーを実行する場合もある。このようなコピーを内部の記憶制御装置が実行している際に、外部の記憶制御装置が、直接接続されている外部のホスト装置からアクセスされると、外部の記憶制御装置は、外部のホスト装置からのアクセスを排除することができないため、該アクセスによって、上記コピーの内容が破壊される不具合が生じる。

【0010】

従って本発明の第1の目的は、或るディスクアレイ装置から見て、外部に存在するデバイス資源である外部のディスクアレイ装置内の記憶デバイスを、上位装置と接続することの可能なデバイスの資源として利用することができるようにすることにある。

【0011】

また、本発明の第2の目的は、或るディスクアレイ装置から見て、外部に存在するホスト装置上のアプリケーションプログラムが、外部のホスト装置からの（リモートコピーの）コマンドに対応不能な外部のディスクアレイ装置からのインクアイリ情報を、対応可能な外部のディスクアレイ装置を通じて取得することができるようにすることにある。

【0012】

更に、本発明の第3の目的は、或るディスクアレイ装置が、該ディスクアレイ装置から見て外部に存在するディスクアレイ装置からのコピーを実行している際に、該外部のディスクアレイ装置が、直接接続されている外部のホスト装置からのアクセスを排除することができるようにすることにある。

【課題を解決するための手段】

【0013】

本発明の第1の観点に従うディスクアレイ装置は、ホストコンピュータ及び／又は上記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるものであって、上記他のディスクアレイ装置の有する記憶領域とこの自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、上記ホストコンピュータから送られてくるデータを記憶するもので、この自ディスクアレイ装置が有する記憶領域と上記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して上記ホストコンピュータからのアクセス対象とする論理ユニット形成部と、上記ホストコンピュータ及び／又は上記他のディスクアレイ装置から転送されるデータを、上記各論理ユニットにライトするデータライト部と、上記ホストコンピュータから自ディスクアレイ装置及び／又は上記他のディスクアレイ装置に格納されているデータのリード要求があった場合に、このデータリード要求の正当性をチェックする第1のチェック部と、上記第1のチェック部が上記データリード要求が正当であると判断した場合に、上記データリード要求に基づいて上記各論理ユニットに保存されているデータを上記ホストコンピュータへ転送するデータ転送部と、上記ホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、この指示の正当性をチェックする第2のチェック部と、上記第2のチェック部による

チェックの結果、上記ホストコンピュータからの上記指示が正当であった場合に、上記指示に応じて上記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成部と、上記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して上記ホストコンピュータからのライトアクセスがあった場合に、この論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするコピー部と、を備え、上記コピー部が、上記複数の論理ユニットについての情報、及び上記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、上記他のディスクアレイ装置が有する実際の記憶領域に対して上記ライトデータをコピーするようにしている。

#### 【0014】

本発明の第1の観点に係る好適な実施形態では、上記データリード要求が、自ディスクアレイ装置に対する上記他のディスクアレイ装置に格納されているデータの取得を指示するためのコマンドと、このコマンドに基づいて自ディスクアレイ装置が上記他のディスクアレイ装置から取得したデータの上記ホストコンピュータへの転送を指示するコマンドと、を含む。

#### 【0015】

上記とは別の実施形態では、上記第1のチェック部が、上記データ取得を指示するためのコマンド、及び上記ホストコンピュータへのデータの転送を指示するコマンドの双方の正当性を夫々チェックする。

#### 【0016】

また、上記とは別の実施形態では、上記第1のチェック部が、上記各コマンドをチェックした結果、双方又は何れか一方のコマンドに正当性が無いと判断した場合には、夫々エラーが生じた旨を上記ホストコンピュータへ報告する。

#### 【0017】

また、上記とは別の実施形態では、上記第1のチェック部が上記データの取得を指示するためのコマンドに正当性があると判断した場合に、上記ホストコンピュータへ転送すべきデータをこのコマンドに基づいて作成するデータ作成部、を更に備える。

#### 【0018】

また、上記とは別の実施形態では、上記データ転送部が、上記データ作成部により作成されたデータを、上記ホストコンピュータが上記データリード要求の送信元であることを確認してから転送するようにしている。

#### 【0019】

更に、上記とは別の実施形態では、上記データライト部が、上記他のディスクアレイ装置から転送されるデータを、上記論理ユニットに接続される仮想化されたデバイスである、自ディスクアレイ装置が有する記憶領域にマッピングすることにより自ディスクアレイ装置へ書き込むようにしている。

#### 【0020】

本発明の第2の観点に従うディスクアレイ装置は、ホストコンピュータ及び／又は上記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるものであって、上記他のディスクアレイ装置の有する記憶領域とこの自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、上記ホストコンピュータから送られてくるデータを記憶するもので、この自ディスクアレイ装置が有する記憶領域と上記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して上記ホストコンピュータからのアクセス対象とする論理ユニット形成部と、自ディスクアレイ装置が有する記憶領域を仮想化した仮想デバイスの識別情報を少なくとも含み、且つ、上記仮想デバイスと上記仮想デバイスにマッピングされた上記他のディスクアレイ装置が有する記憶領域との対応関係を示すマッピングテーブルを保持するマッピングテーブル保持部と、上記ホストコンピュータから少なくとも上記仮想デバイスの識別情報を含

むデータリード要求が送信された場合に、上記識別情報に基づいて対象となる仮想デバイスを上記マッピングテーブルから検索する検索部と、上記検索部によって検索された仮想デバイスからデータをリードして上記ホストコンピュータへ転送するデータ転送部と上記ホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、この指示の正当性をチェックするチェック部と、上記チェック部によるチェックの結果、上記ホストコンピュータからの上記指示が正当であった場合に、上記指示に応じて上記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成部と、上記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して上記ホストコンピュータからのライトアクセスがあった場合に、この論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするコピー部と、を備え、上記コピー部が、上記複数の論理ユニットについての情報、及び上記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、上記他のディスクアレイ装置が有する実際の記憶領域に対して上記ライトデータをコピーするようにしている。

#### 【0021】

本発明の第2の観点に係る好適な実施形態では、上記仮想デバイスには、自ディスクアレイ装置が有する別の記憶領域のデータもコピーされるようになっており、上記ホストコンピュータからのデータリード要求には、上記別の記憶領域の識別情報も含まれる。

#### 【0022】

上記とは別の実施形態では、上記データ転送部によって自ディスクアレイ装置から上記ホストコンピュータに転送されたデータが、上記ホストコンピュータから自ディスクアレイ装置には直接に接続されることが無い別のホストコンピュータへ更に転送される。

#### 【0023】

本発明の第3の観点に従うディスクアレイ装置は、ホストコンピュータ及び／又は上記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるものであって、上記他のディスクアレイ装置の有する記憶領域とこの自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、上記ホストコンピュータから送られてくるデータを記憶するもので、この自ディスクアレイ装置が有する記憶領域と上記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して上記ホストコンピュータからのアクセス対象とする論理ユニット形成部と、上記ホストコンピュータ及び／又は上記他のディスクアレイ装置から転送されるデータを、上記各論理ユニットにライトするデータライト部と、上記ホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、この指示の正当性をチェックするチェック部と、上記チェック部によるチェックの結果、上記ホストコンピュータからの上記指示が正当であった場合に、上記指示に応じて上記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成部と、上記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して上記ホストコンピュータからのライトアクセスがあった場合に、この論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするコピー部と、を備え、上記コピー部が、上記複数の論理ユニットについての情報、及び上記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、上記他のディスクアレイ装置が有する実際の記憶領域に対して上記ライトデータをコピーするようにしている。

## 【0024】

本発明の第3の観点に係る好適な実施形態では、上記チェック部が、上記ペア作成の指示の正当性をチェックした結果、正当性が無いと判断した場合には、エラーが生じた旨を前記ホストコンピュータへ報告するようにしている。

## 【0025】

上記とは別の実施形態では、上記ペア状態形成部が、上記ペア状態形成の処理において問題が生じたと判断した場合には、エラーが生じた旨を上記ホストコンピュータへ報告するようにしている。

## 【0026】

本発明の第4の観点に従うディスクアレイ装置は、第1のホストコンピュータ及び／又は第2のホストコンピュータに接続された他のディスクアレイ装置と通信可能に接続されるものであって、上記他のディスクアレイ装置が有する記憶領域とこの自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、上記第1のホストコンピュータから送られてくるデータを記憶するもので、この自ディスクアレイ装置が有する記憶領域と上記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して上記ホストコンピュータからのアクセス対象とする論理ユニット形成部と、上記ホストコンピュータ及び／又は上記他のディスクアレイ装置から転送されるデータを、上記各論理ユニットにライトするデータライト部と、上記第1のホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、この指示の正当性をチェックするチェック部と、上記チェック部によるチェックの結果、上記第1のホストコンピュータからの上記指示が正当であった場合に、上記指示に応じて上記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成部と、上記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して上記第1のホストコンピュータからのライトアクセスがあった場合に、この論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするコピー部と、上記第2のホストコンピュータから自ディスクアレイ装置及び／又は上記他のディスクアレイ装置に格納されているデータのリード要求があった場合に、上記データリード要求に基づく処理が開始されてから終了するまでの間、上記第2のホストコンピュータによる上記他のディスクアレイ装置へのアクセスを禁止する禁止部と、を備え、上記コピー部が、上記複数の論理ユニットについての情報、及び上記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、上記他のディスクアレイ装置が有する実際の記憶領域に対して上記ライトデータをコピーするようにしている。

## 【0027】

本発明の第4の観点に係る好適な実施形態では、上記アクセス禁止が、上記第2のホストコンピュータによる上記他のディスクアレイ装置へのアクセスから上記他のディスクアレイ装置のデータを保護するための上記禁止部による上記他のディスクアレイ装置へのリザーブコマンドの発行である。

## 【0028】

上記とは別の実施形態では、上記仮想デバイスには、自ディスクアレイ装置が有する別の記憶領域のデータもコピーされるようになっている。

## 【0029】

本発明の第5の観点に従うディスクアレイ装置の制御方法は、ホストコンピュータ及び／又は上記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、上記他のディスクアレイ装置の有する記憶領域とこの自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、上記ホストコンピュータから送られてくるデータを記憶するディスクアレイ装置を対象とするもの



で、この自ディスクアレ装置が有する記憶領域と上記他のディスクアレ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して上記ホストコンピュータからのアクセス対象とする論理ユニット形成のステップと、上記ホストコンピュータ及び／又は上記他のディスクアレ装置から転送されるデータを、上記各論理ユニットにライトするデータライトのステップと、上記ホストコンピュータから自ディスクアレ装置及び／又は上記他のディスクアレ装置に格納されているデータのリード要求があった場合に、このデータリード要求の正当性をチェックする第1のチェックのステップと、上記第1のチェックのステップにおいて上記データリード要求が正当であると判断した場合に、上記データリード要求に基づいて上記各論理ユニットに保存されているデータを上記ホストコンピュータへ転送するデータ転送のステップと、上記ホストコンピュータから、自ディスクアレ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、この指示の正当性をチェックする第2のチェックのステップと、上記第2のチェックのステップにおけるチェックの結果、上記ホストコンピュータからの上記指示が正当であった場合に、上記指示に応じて上記複数の論理ユニットのうち、自ディスクアレ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成のステップと、上記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレ装置が有する記憶領域から形成された論理ユニットに対して上記ホストコンピュータからのライトアクセスがあった場合に、この論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするライトデータコピーのステップと、を備え、上記ライトデータコピーのステップが、上記複数の論理ユニットについての情報、及び上記他のディスクアレ装置が有する記憶領域と自ディスクアレ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、上記他のディスクアレ装置が有する実際の記憶領域に対して上記ライトデータをコピーするようにしている。

### 【0030】

本発明の第6の観点に従うディスクアレ装置の制御方法は、ホストコンピュータ及び／又は上記ホストコンピュータとは直接接続できない他のディスクアレ装置と通信可能に接続されるディスクアレ装置であって、上記他のディスクアレ装置の有する記憶領域とこの自ディスクアレ装置が有する記憶領域とを併せた記憶領域に対して、上記ホストコンピュータから送られてくるデータを記憶するディスクアレ装置を対象とするもので、この自ディスクアレ装置が有する記憶領域と上記他のディスクアレ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して上記ホストコンピュータからのアクセス対象とする論理ユニット形成のステップと、自ディスクアレ装置が有する記憶領域を仮想化した仮想デバイスの識別情報を少なくとも含み、且つ、上記仮想デバイスと上記仮想デバイスにマッピングされた上記他のディスクアレ装置が有する記憶領域との対応関係を示すマッピングテーブルを保持するマッピングテーブル保持のステップと、上記ホストコンピュータから少なくとも上記仮想デバイスの識別情報を含むデータリード要求が送信された場合に、上記識別情報に基づいて対象となる仮想デバイスを上記マッピングテーブルから検索する検索のステップと、上記検索のステップにおいて検索された仮想デバイスからデータをリードして上記ホストコンピュータへ転送するデータ転送のステップと、上記ホストコンピュータから、自ディスクアレ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、この指示の正当性をチェックするチェックのステップと、上記チェックのステップにおけるチェックの結果、上記ホストコンピュータからの上記指示が正当であった場合に、上記指示に応じて上記複数の論理ユニットのうち、自ディスクアレ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成のステップと、上記ペア状態を形成している



2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して上記ホストコンピュータからのライトアクセスがあった場合に、この論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするライトデータコピーのステップと、を備え、上記ライトデータコピーのステップが、上記複数の論理ユニットについての情報、及び上記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、上記他のディスクアレイ装置が有する実際の記憶領域に対して上記ライトデータをコピーするようにしている。

#### 【0031】

本発明の第7の観点に従うディスクアレイ装置の制御方法は、ホストコンピュータ及び／又は上記ホストコンピュータとは直接接続できない他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、上記他のディスクアレイ装置の有する記憶領域とこの自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、上記ホストコンピュータから送られてくるデータを記憶するディスクアレイ装置を対象とするもので、この自ディスクアレイ装置が有する記憶領域と上記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して上記ホストコンピュータからのアクセス対象とする論理ユニット形成のステップと、上記ホストコンピュータ及び／又は上記他のディスクアレイ装置から転送されるデータを、上記各論理ユニットにライトするデータライトのステップと、上記ホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、この指示の正当性をチェックするチェックのステップと、上記チェックのステップにおけるチェックの結果、上記ホストコンピュータからの上記指示が正当であった場合に、上記指示に応じて上記複数の論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成のステップと、上記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して上記ホストコンピュータからのライトアクセスがあった場合に、この論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするライトデータコピーのステップと、を備え、上記ライトデータコピーのステップが、上記複数の論理ユニットについての情報、及び上記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、上記他のディスクアレイ装置が有する実際の記憶領域に対して上記ライトデータをコピーするようにしている。

#### 【0032】

本発明の第8の観点に従うディスクアレイ装置の制御方法は、第1のホストコンピュータ及び／又は第2のホストコンピュータに接続された他のディスクアレイ装置と通信可能に接続されるディスクアレイ装置であって、上記他のディスクアレイ装置が有する記憶領域とこの自ディスクアレイ装置が有する記憶領域とを併せた記憶領域に対して、上記第1のホストコンピュータから送られてくるデータを記憶するディスクアレイ装置を対象とするもので、この自ディスクアレイ装置が有する記憶領域と上記他のディスクアレイ装置が有する記憶領域とを併せた記憶領域から複数の論理ユニットを形成して上記ホストコンピュータからのアクセス対象とする論理ユニット形成のステップと、上記ホストコンピュータ及び／又は上記他のディスクアレイ装置から転送されるデータを、上記各論理ユニットにライトするデータライトのステップと、上記第1のホストコンピュータから、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成すべき旨の指示が送信された場合に、この指示の正当性をチェックするチェックのステップと、上記チェックのステップにおけるチェックの結果、上記第1のホストコンピュータからの上記指示が正当であった場合に、上記指示に応じて上記複数の論理ユニットのうち、自ディスクア

レイ装置が有する記憶領域から形成された論理ユニットと上記他のディスクアレイ装置が有する記憶領域から形成された論理ユニットとの間でペア状態を形成するペア状態形成のステップと、上記ペア状態を形成している2つの論理ユニットのうち、自ディスクアレイ装置が有する記憶領域から形成された論理ユニットに対して上記第1のホストコンピュータからのライトアクセスがあった場合に、この論理ユニットとペア状態を形成している他方の論理ユニットに対してライトデータをコピーするライトデータコピーのステップと、上記第2のホストコンピュータから自ディスクアレイ装置及び／又は上記他のディスクアレイ装置に格納されているデータのリード要求があった場合に、上記データリード要求に基づく処理が開始されてから終了するまでの間、上記第2のホストコンピュータによる上記他のディスクアレイ装置へのアクセスを禁止するアクセス禁止のステップと、を備え、上記ライトデータコピーのステップが、上記複数の論理ユニットについての情報、及び上記他のディスクアレイ装置が有する記憶領域と自ディスクアレイ装置が有する記憶領域とを併せた記憶領域についての情報との関係を利用して、上記他のディスクアレイ装置が有する実際の記憶領域に対して上記ライトデータをコピーするようにしている。

#### 【発明の効果】

##### 【0033】

本発明によれば、或るディスクアレイ装置から見て、外部に存在するデバイス資源である外部のディスクアレイ装置内の記憶デバイスを、上位装置と接続することの可能なデバイスの資源として利用することができるようにすることができる。

#### 【発明を実施するための最良の形態】

##### 【0034】

以下、本発明の実施の形態を、図面により詳細に説明する。

##### 【0035】

本発明に係る実施形態では、以下に詳細に説明するように、自己（即ち、自記憶制御装置）の外部（即ち、他の記憶制御装置内）に存在する記憶デバイスを自己の仮想デバイス（VDEV）にマッピングすることにより、自己の外部に存在する記憶デバイスを自己の内部ボリュームとして、上位装置（例えば、パーソナルコンピュータ、メインフレーム等のコンピュータ）に提供するようになっている。

##### 【0036】

図1は、本発明の第1の実施形態に係る記憶システムの要部の構成を示すブロック図である。

##### 【0037】

図1において、ホスト装置10は、例えば、CPU（Central Processing Unit）やメモリ等の情報処理資源を備えるコンピュータ装置であり、例えば、パーソナルコンピュータ、ワークステーション、メインフレーム等がホスト装置10として用いられる。ホスト装置10は、例えば、キーボードスイッチや、ポインティングデバイスや、マイクロフォン等の情報入力装置（図示しない）と、例えば、モニタディスプレイや、スピーカ等の情報出力装置（図示しない）とを備える。ホスト装置10は、上記各部に加えて、更に、例えば、第1の記憶制御装置20が提供する記憶領域を使用するデータベースソフトウェア等のアプリケーションプログラム11と、通信ネットワークCN1を介して第1の記憶制御装置20にアクセスするためのアダプタ（HBA）12とを備える。

##### 【0038】

ホスト装置10は、通信ネットワークCN1を介して第1の記憶制御装置20に接続されるが、通信ネットワークCN1としては、例えば、LAN（Local Area Network）、SAN（Storage Area Network）、インターネット、専用回線、公衆（電話）回線等の場合に応じて適宜用いることができる。LANを介するデータ通信は、例えば、TCP/IP（Transmission Control Protocol/Internet

Protocol）プロトコルに従って行われる。通信ネットワークCN1としてLANが用いられる場合、ホスト装置10は、第1の記憶制御装置20に対し、ファイル名を指定してフ

ファイル単位でのデータ入出力を要求する。通信ネットワークCN1としてSANが用いられる場合、ホスト装置10は、第1の記憶制御装置20に対し、ファイバチャネルプロトコルに従って、複数のディスク記憶装置（ディスクドライブ）により提供される記憶領域のデータ管理単位であるブロックを単位としてデータ入出力を要求する。アダプタ12は、LANが通信ネットワークCN1として用いられる場合には、例えばLAN対応のネットワークカードであり、SANが通信ネットワークCN1として用いられる場合には、例えばホストバスアダプタである。

#### 【0039】

なお、図1では、ホスト装置10は、通信ネットワークCN1を介して第1の記憶制御装置20にのみ接続されるように記載されているが、通信ネットワークCN2を介してホスト装置10を第2の記憶制御装置40にも接続するようにしてもよい。また、第2の通信ネットワークCN2としては、例えば第1の通信ネットワークCN1と同様に、SAN、LAN、インターネット、専用回線、公衆（電話）回線等の場合に応じて適宜用いることができる。

#### 【0040】

第1の記憶制御装置20は、例えば、ディスクアレイサブシステムや、高機能化されたインテリジェント型のファイバチャネルスイッチ等として構成され得るものであり、後に説明するように、第2の記憶制御装置40が有する記憶資源を自己（即ち、第1の記憶制御装置20）の論理ボリューム（Logical Unit）としてホスト装置10に提供するものである。よって、第1の記憶制御装置20は、自己（即ち、第1の記憶制御装置20）が直接支配するローカルな記憶デバイスを有している必要はない。

#### 【0041】

第1の記憶制御装置20の内部構成は、コントローラ部と記憶装置部とに大別される。コントローラ部は、例えば、複数のチャネルアダプタ（CHA）21と、複数のディスクアダプタ（DKA）22と、コントロールユニット（CU）23と、キャッシュメモリ24と、共有メモリ25と、接続部26とを備える。一方、記憶装置部、即ち、記憶装置30は、複数の記憶デバイス31（32）を含む。

#### 【0042】

コントロール部において、各チャネルアダプタ21は、ホスト装置10との間のデータ通信を行うもので、夫々がホスト装置10との間で通信を行うための通信ポート21Aを備える。各チャネルアダプタ21は、夫々CPUやメモリ等を有するマイクロコンピュータシステムとして構成されており、ホスト装置10から受信した各種コマンドを解釈して実行する。また、各チャネルアダプタ21には、夫々を識別するためのネットワークアドレス（例えば、IPアドレスやWWN等）が割り当てられており、各チャネルアダプタ21が、夫々個別にNAS（Network Attached Storage）として振る舞えるようになっている。そのため、ホスト装置10が複数存在する場合には、各チャネルアダプタ21が、各ホスト装置10からの要求を夫々個別に受け付けることが可能である。

#### 【0043】

各ディスクアダプタ22は、記憶装置30が含む記憶デバイス31、32との間のデータ授受を行うもので、夫々が記憶デバイス31、32に接続するための通信ポート22Aを備える。各ディスクアダプタ22は、夫々CPUやメモリ等を有するマイクロコンピュータシステムとして構成されており、チャネルアダプタ21がホスト装置10から受信したデータを、接続部26を介して読み込み、該データをホスト装置10からのリクエスト（書き込み命令）に基づいて、所定の記憶デバイス31、32の所定のアドレスに書込む。各ディスクアダプタ22は、また、チャネルアダプタ21、及び接続部26を介して与えられるホスト装置10からのリクエスト（読み出し命令）に基づいて、所定の記憶デバイス31、32の所定のアドレスからデータを読み出し、該データを、接続部26、及びチャネルアダプタ21を介してホスト装置10に送信する。なお、各ディスクアダプタ22は、記憶デバイス31、32との間でデータ入/出力を行う場合には、論理的なアドレスを物理的なアドレスに変換する。また、各ディスクアダプタ22は、記憶デバイス31、

32がRAIDに従って管理されている場合には、RAID構成に応じたデータアクセスを行う。

【0044】

コントロールユニット23は、第1の記憶制御装置20全体の動作を制御するもので、例えば、管理用のコンソール（図示しない）が接続されている。コントロールユニット23は、第1の記憶制御装置20内の障害発生を監視して監視結果を示す情報を該コンソール（図示しない）に表示したり、コンソール（図示しない）からの指令に基づいて、記憶ディスク（例えば、記憶デバイス31、32）の閉塞処理等を指示したりするようになっている。

【0045】

キャッシュメモリ24は、接続部26を介して各チャネルアダプタ21から与えられる、各チャネルアダプタ21がホスト装置10から受信したデータを一時的に記憶すると共に、接続部26を介して各ディスクアダプタ22から与えられる、各ディスクアダプタ22が記憶デバイス31、32から読み出したデータを一時的に記憶する。なお、キャッシュメモリ24に代えて、記憶デバイス31、32のいずれか1つ或いは複数、キャッシュ用のディスク（キャッシュメモリ）として使用しても差し支えない。

【0046】

共有メモリ25には、制御情報等が格納され、また、後述するマッピングテーブルTm等の各種テーブル類も格納されるほか、ワーク領域が設定される。

【0047】

接続部26は、各チャネルアダプタ21、各ディスクアダプタ22、コントロールユニット23、キャッシュメモリ24、及び共有メモリ25の間を相互に接続するもので、接続部26は、例えば、高速スイッチング動作によってデータ伝送を行う超高速クロスバススイッチ等のような高速バスとして構成することが可能である。

【0048】

記憶装置部、即ち、記憶装置30は、複数の記憶デバイス31を備える。記憶デバイス31としては、例えば、ハードディスク、フレキシブルディスク、磁気テープ、半導体メモリ、光ディスク等のような各種のデバイスを用いることができる。記憶装置30において、破線で記載する記憶デバイス32は、第2の記憶制御装置40が有する記憶デバイス42を、第1の記憶制御装置20側に取り込んだ状態を示している。

【0049】

即ち、本実施形態では、第1の記憶制御装置20が、自己（即ち、第1の記憶制御装置20）から見て外部に存在する（第2の記憶制御装置40が有する）記憶デバイス42を、自己の内部記憶デバイスとしてホスト装置10に認識させることによって、ホスト装置10に外部に存在する記憶デバイス42の記憶資源を提供する。

【0050】

第2の記憶制御装置40は、通信ポート41と記憶デバイス42とを備えている。この他にも、第2の記憶制御装置40は、チャネルアダプタやディスクアダプタ等を備えることもできる。しかし、第2の記憶制御装置40の詳細な構成は、本発明の要旨に直接関連しないので、それらの説明を省略する。第2の記憶制御装置40は、通信ネットワークCN2を介して第1の記憶制御装置20に接続されており、第2の記憶制御装置40の記憶デバイス42は、第1の記憶制御装置20の内部記憶デバイスとして扱われるようになっている。

【0051】

図2は、図1で示した第1の記憶制御装置20、及び記憶デバイス32の1つの論理的な概略構造を示す模式図である。

【0052】

図2に示すように、第1の記憶制御装置20は、下層側から順番に、VDEV101と、LDEV102と、LUN103とから成る3層の（論理的な）記憶階層を有している。

【0053】

VDEV101は、論理的な記憶階層の最下位に位置する仮想デバイス (Virtual Device) であって、物理的な記憶資源を仮想化したものであり、RAID構成を適用することができる。即ち、1つの記憶デバイス31から複数のVDEV101を形成することもできるし (これを、スライシングという)、複数の記憶デバイス31から1つのVDEV101を形成することもできる (これを、ストライピングという)。図2に記載した2つのVDEV101のうちの、左側に位置するVDEV101は、例えば、所定のRAID構成に従って記憶デバイス31を仮想化したものである。

#### 【0054】

一方、図2に記載した右側に位置するVDEV101は、第2の記憶制御装置40が有する記憶デバイス42をマッピングすることにより構成されている。即ち、本実施形態では、第2の記憶制御装置40が有する記憶デバイス42により提供される論理ボリューム (LDEV) を、後に説明するマッピングテーブルTmを用いてVDEV101にマッピングすることにより、第1の記憶制御装置20の内部ボリュームとして使用できるようになっている。図2に記載した例では、4つの記憶デバイス42A~42Dをストライピングすることにより、図2に記載した右側に位置するVDEV101が構築されている。各記憶デバイス42A~42Dには、夫々の通信ポート41A~41Dから夫々のLUN43A~43Dを特定することにより、夫々個別にアクセスすることが可能である。各通信ポート41A~41Dには、ユニークな識別情報であるWWNが割り当てられており、また、各LUN43A~43Dには、LUN番号が設定されているので、WWN及びLUN番号の組合せによって記憶デバイスを特定することが可能である。

#### 【0055】

VDEV101の (論理的に) 1層上 (の記憶階層) には、LDEV102が設定される。LDEV102は、仮想デバイス (VDEV) を更に仮想化した論理デバイス (論理ボリューム) である。1つのVDEV101から2つのLDEV102に接続することも可能であるし、複数のVDEV101から1つのLDEV102に接続することも可能である。ホスト装置10は、LDEV102に対し、夫々のLUN103を介してアクセスすることが可能である。このように、本実施形態では、LUN103と記憶デバイス42との間に位置する中間記憶階層 (VDEV101, LDEV102) に記憶デバイス42を接続することにより、第1の記憶制御装置20から見て外部に存在する (第2の記憶制御装置40が有する) 記憶デバイス42を、第1の記憶制御装置20の内部ボリュームの1つとして利用できるようにしている。

#### 【0056】

図3は、図1で示した第1の記憶制御装置20、及び記憶デバイス32の他の論理的な概略構造を示す模式図である。

#### 【0057】

図3において、LDEV50は、第2の記憶制御装置40が有する複数の記憶デバイス42により提供されるもので、LDEV50は、複数の経路を備える交代パス構成を有している。即ち、各記憶デバイス42の上には、図示のように、論理ボリュームであるLDEV50が構築されており、このLDEV50には、第1の記憶制御装置20側から2つの経路 (アクセスデータパス) を介して夫々アクセスすることが可能である。一方の経路は、第1の通信ポート41 (1) からLUN43を介してLDEV50に到達するものであり、他方の経路は、第2の通信ポート41 (2) から上記とは別のLUN43を介してLDEV50に到達するものである。従って、仮に上記2つの経路のうちのいずれか一方が、障害等で使用不能な場合であっても、第1の記憶制御装置20側から他方の経路を介してLDEV50にアクセスすることができる。

#### 【0058】

図3で示したように、第1の記憶制御装置20側から複数の経路を介してLDEV50に夫々アクセスが可能な場合において、第1の記憶制御装置20側から一方の経路を介してLDEV50に格納されているデータを利用中に、他方の経路を介してLDEV50にアクセスして第1の記憶制御装置20側から上記データの更新等を行うことが無いように

、必要なデータ保護等が施される。

#### 【0059】

なお、図3に記載した例では、第1の記憶制御装置20は、第2の記憶制御装置40の記憶資源(LDEV50)を自己のVDEV101にマッピングすることにより、外部のLDEV50を内部のLDEV102として利用している。また、複数のLDEV102を1つのVDEV101上に設定し、このVDEV101には外部のLDEV50が複数の経路を介してマッピングされている。

#### 【0060】

ホスト装置10は、(第1の記憶制御装置20の)LUN103のみを認識しており(結果的にLDEV102まで認識しており)、LUN103よりも階層が下の構造については、ホスト装置10に対して隠されている。第1の記憶制御装置20において、複数のLDEV102は、夫々同一のVDEV101を利用しており、このVDEV101は、複数の経路を介して同一のLDEV50に接続されている。従って、図3に示す例では、第2の記憶制御装置40の有する交代パス構造を利用して、第1の記憶制御装置20の冗長性を高めることが可能である。

#### 【0061】

図4は、第1の記憶制御装置20から見て外部に存在する記憶デバイス42(詳しくは、第2の記憶制御装置40が有する複数の記憶デバイス42により提供されるLDEV50)を、(第1の記憶制御装置20の)VDEV101にマッピングするためのテーブル構造の一例を示す説明図である。

#### 【0062】

図4において、マッピングテーブルT<sub>m</sub>は、例えば、VDEV101を夫々識別するためのVDEV番号と、外部の記憶デバイス42の情報とを夫々対応付けた構成になっている。外部デバイス情報としては、例えば、デバイス識別情報と、各記憶デバイス42の記憶容量と、デバイスの種別を示す情報(例えば、テープ系デバイスかディスク系デバイスか等)と、各記憶デバイス42へのパス情報とを含んだ構成となっている。また、パス情報は、(図1乃至図3で示した)各通信ポート41に固有の識別情報(WWN)と、LUN43を識別するためのLUN番号とを含んだ構成となっている。

#### 【0063】

なお、図4に記載されているデバイス識別情報やWWN等は、説明の便宜のために設定された値であって、特に意味は無い。また、図4の下側に記載されているVDEV番号「3」のVDEV101には、図示のように、3個のパス情報が対応付けられている。換言すれば、このVDEV101(#3)にマッピングされる外部記憶デバイス42は、その内部に3つの経路を有する交代パス構造を備えており、VDEV101(#3)には、この交代パス構造を認識して、上記外部記憶デバイス42がマッピングされている。上記3つの経路のいずれを通過しても、第1の記憶制御装置20側から(外部記憶デバイス42における)同一の記憶領域にアクセスできることが判明しているため、いずれか1つ又は2つの経路に障害等が発生した場合でも、残りの正常な経路を介して(第1の記憶制御装置20側から外部記憶デバイス42内の)所望のデータにアクセスすることが可能である。

#### 【0064】

図4に記載したようなマッピングテーブルT<sub>m</sub>を採用することにより、第1の記憶制御装置20内の1つ以上のVDEV101に対し、1つ又は複数の外部の記憶デバイス42をマッピングすることが可能である。

#### 【0065】

次に、図5を参照して、(第1の記憶制御装置20から見て)外部の記憶デバイス42をVDEV101にマッピングする方法の一例を説明する。

#### 【0066】

図5は、マッピング時に、第1の記憶制御装置20と第2の記憶制御装置40との間で行われる処理動作の要部を示すタイムチャートである。

#### 【0067】

図5において、まず、第1の記憶制御装置20は、チャネルアダプタ21のイニシエータポート(21A)を介して、第2の記憶制御装置40にログインする(ステップS1)。この処理動作の後に、第2の記憶制御装置40が、第1の記憶制御装置20のログインに対して応答を返すことにより、ログインが完了する(ステップS2)。次に、第1の記憶制御装置20は、例えば、SCSI (Small Computer System Interface) 規格で定められている照会コマンド(inquiryコマンド)を、第2の記憶制御装置40に送信して、第2の記憶制御装置40の有する記憶デバイス42の詳細について応答を求める(ステップS3)。

#### 【0068】

ここで、照会コマンドは、照会先の装置の種類及び構成を明らかにするために用いられるもので、照会先装置の有する(記憶)階層を透過してその物理的構造を把握することができる。照会コマンドを使用することにより、第1の記憶制御装置20は、例えば、装置名、デバイスタイプ、製造番号(プロダクトID)、LDEV番号、各種バージョン情報、ベンダID等の情報を第2の記憶制御装置40から取得することができる(ステップS4)。第2の記憶制御装置40は、問合せされた情報を第1の記憶制御装置20に送信し、応答する(ステップS5)。

#### 【0069】

次に、第1の記憶制御装置20は、第2の記憶制御装置40から取得した情報を、(図4で示した)マッピングテーブルTmの所定箇所に登録した後(ステップS6)、第2の記憶制御装置40から記憶デバイス42の記憶容量を読み出す(ステップS7)。第2の記憶制御装置40は、第1の記憶制御装置20からの問合せに対して、記憶デバイス42の記憶容量を返信し(ステップS8)、応答を返す(ステップS9)。第1の記憶制御装置20は、記憶デバイス42の記憶容量を(図4で示した)マッピングテーブルTmの所定箇所に登録する(ステップS10)。

#### 【0070】

以上の処理動作を実行することにより、(図4で示した)マッピングテーブルTmを構築することができる。第1の記憶制御装置20のVDEV101にマッピングされた外部の記憶デバイス42(外部LUN、即ち外部のLDEV50)との間でデータの入/出力を行う場合には、後に説明する他のテーブルを参照してアドレス変換等を行うことになる。

#### 【0071】

次に、図6乃至図8を参照して、第1の記憶制御装置20と第2の記憶制御装置40との間のデータ入/出力について説明する。まず最初に、データを書き込む場合について、図6及び図7に基づいて説明する。

#### 【0072】

図6は、ホスト装置10による第1の記憶制御装置20の論理ボリュームに対するデータ書き込み時の処理を示す模式図であり、図7は、図6に記載した処理動作の流れを各種テーブルとの関係で示す説明図である。

#### 【0073】

ホスト装置10は、第1の記憶制御装置20が提供する論理ボリューム(LDEV102)にデータを書き込むことができる。例えば、SANの中に仮想的なSANサブネットを設定するゾーニングや、アクセス可能なLUNのリストをホスト装置10が保持するLUNマスキングという手法により、ホスト装置10を特定のLDEV102に対してのみアクセスさせるように設定することができる。

#### 【0074】

ホスト装置10がデータを書き込もうとするLDEV102が、VDEV101を介して第1の記憶制御装置20内部の記憶デバイスである記憶デバイス31に接続されている場合、通常の処理によってデータが書き込まれる。即ち、ホスト装置10からのデータは、一旦キャッシュメモリ24に格納され、キャッシュメモリ24からディスクアダプタ22を介して、所定の記憶デバイス31の所定アドレスに格納される。この際、ディスクア



データ 22 は、論理的なアドレスを物理的なアドレスに変換する。また、RAID 構成の場合、同一のデータが複数の記憶デバイス 31 に記憶等される。

#### 【0075】

これに対し、ホスト装置 10 が書き込もうとする LDEV 102 が、VDEV 102 を介して（第 1 の記憶制御装置 20 から見て）外部の記憶デバイス 42 に接続されている場合、図 6 に示すような処理の流れでデータが書き込まれる。図 6 (a) は記憶階層を中心に示す流れ図であり、図 6 (b) はキャッシュメモリ 24 の使われ方を中心に示す流れ図である。

#### 【0076】

図 6 において、まずホスト装置 10 は、書き込み先の LDEV 102 を特定する LDEV 番号と、この LDEV 102 にアクセスするための通信ポート 21A を特定する WWN とを明示して、書き込みコマンド (Write) を第 1 の記憶制御装置 20 に対して発行する（ステップ S21）。この書き込みコマンドを受信すると、第 1 の記憶制御装置 20 は、第 2 の記憶制御装置 40 に送信するための書き込みコマンドを生成して、該書き込みコマンドを、第 2 の記憶制御装置 40 に送信する（ステップ S22）。第 1 の記憶制御装置 20 は、ホスト装置 10 から受信した書き込みコマンド中の書き込み先アドレス情報等を、外部 LDEV 50 に合わせて変更することによって、新たな書き込みコマンドを生成する。

#### 【0077】

次に、ホスト装置 10 は、書き込むべきデータを第 1 の記憶制御装置 20 に送信する（ステップ S23）と、第 1 の記憶制御装置 20 に受信された（上記書き込むべき）データは、LDEV 102 から VDEV 101 を介して（ステップ S24）、外部の LDEV 50 に転送される（ステップ S26）。ここで、第 1 の記憶制御装置 20 は、ホスト装置 10 からのデータをキャッシュメモリ 24 に格納した時点で、ホスト装置 10 に対して書き込み完了の応答 (Good) を返す（ステップ S25）。第 2 の記憶制御装置 40 は、第 1 の記憶制御装置 20 からデータを受信した時点で（或いは記憶デバイス 42 に書き込みを終えた時点で）、書き込み完了報告を第 1 の記憶制御装置 20 に送信する（ステップ S26）。即ち、第 1 の記憶制御装置 20 がホスト装置 10 に対して書き込み完了を報告する時期（ステップ S25）と、実際にデータが記憶デバイス 42 に記憶される時期とは相違する（非同期方式）。従って、ホスト装置 10 は、実際にデータが記憶デバイス 42 に格納される前にデータ書き込み処理から解放され、別の処理を行うことができる。

#### 【0078】

図 6 (b) において、キャッシュメモリ 24 には、多数のサブブロック 24A が設けられている。第 1 の記憶制御装置 20 は、ホスト装置 10 から指定された論理ブロックアドレスをサブブロックのアドレスに変換し、キャッシュメモリ 24 の所定箇所にデータを格納する（ステップ S24）。

#### 【0079】

次に、図 7 を参照し、各種テーブルを利用してデータが変換される様子を説明する。図 7 の上部に示すように、ホスト装置 10 は、所定の通信ポート 21A に対し、LUN 番号 (LUN#)、及び論理ブロックアドレス (LBA) を指定してデータを送信する。第 1 の記憶制御装置 20 は、LDEV 102 用に入力されたデータ (LUN#+LBA) を、図 7 (a) に示す第 1 の変換テーブル T1 に基づいて、VDEV 101 用のデータに変換する。第 1 の変換テーブル T1 は、内部の LUN103 を指定するデータを VDEV 101 用データに変換するための、LUN-LDEV-VDEV 変換テーブルである。

#### 【0080】

このテーブル T1 は、例えば、LUN 番号 (LUN#) と、その LUN103 に対応する LDEV 102 の番号 (LDEV#)、及び最大スロット数と、LDEV 102 に対応する VDEV 101 の番号 (VDEV#)、及び最大スロット数等を対応付けることにより構成される。このテーブル T1 を参照することにより、ホスト装置 10 からのデータ (LUN#+LBA) は、VDEV 101 用のデータ (VDEV#+SLOT#+SUBBLOCK#) に変換される。



## 【0081】

次に、第1の記憶制御装置20は、図7(b)に示す第2の変換テーブルT2を参照して、VDEV101用のデータを、第2の記憶制御装置40の外部LUN(LDEV50)用に送信して記憶させるためのデータに変換する。第2の変換テーブルT2には、例えば、VDEV101の番号(VDEV#)と、そのVDEV101からのデータを第2の記憶制御装置40に送信するためのイニシエータポートの番号と、データ転送先の通信ポート41を特定するためのWWNと、その通信ポートを介してアクセス可能なLUN番号とが対応付けられている。この第2の変換テーブルT2に基づいて、第1の記憶制御装置20は、記憶させるべきデータの宛先情報を、イニシエータポート番号# + WWN + LUN# + LBAの形式に変換する。このように宛先情報が変更されたデータは、指定されたイニシエータポートから通信ネットワークCN2を介して、指定された通信ポート41に到達する。そして、データは、指定されたLUN43でアクセス可能なLDEV50の所定の場所に格納される。既に説明したように、LDEV50は、複数の記憶デバイス42上に仮想的に構築されているので、データのアドレスは物理アドレスに変換されて、所定のディスクの所定アドレスに格納される。

## 【0082】

図7(c)は、別の第2の変換テーブルT2aを示す。この変換テーブルT2aは、外部記憶デバイス42に由来するVDEV101に、ストライプやRAIDを適用する場合に使用される。変換テーブルT2aは、VDEV番号(VDEV#)と、ストライプサイズと、RAIDレベルと、第2の記憶制御装置40を識別するための番号(SS#(ストレージシステム番号))と、イニシエータポート番号と、通信ポート41のWWN及びLUN43の番号とを対応付けることにより構成される。図7(c)に示す例では、1つのVDEV101は、SS#(1、4、6、7)で特定される合計4つの外部記憶制御装置を利用してRAID1を構成する。また、SS#1に割り当てられている3個のLUN(#0、#0、#4)は、同一デバイス(LDEV#)に設定されている。なお、LUN#0のボリュームは、2個のアクセスデータパスを有する交代パス構造を備える。このように、本実施形態では、外部に存在する複数の論理ボリューム(LDEV)からVDEV101を構成することにより、ストライピングやRAID等の機能を追加した上でホスト装置10に提供することができる。

## 【0083】

次に、図8を参照して、ホスト装置10が、第2の記憶制御装置40のLDEV50からデータを読み出す場合の処理の流れを説明する。

## 【0084】

図8において、まず、ホスト装置10は、通信ポート21Aを指定して第1の記憶制御装置20にデータの読み出しコマンドを送信する(ステップS31)。この読み出しコマンドを受信すると、第1の記憶制御装置20は、要求されたデータを第2の記憶制御装置40から読み出すべく、読み出しコマンドを生成し、該読み出しコマンドを第2の記憶制御装置40に送信する(ステップS32)。この読み出しコマンドを受信すると、第2の記憶制御装置40は、該読み出しコマンドに応じて、要求されたデータを記憶デバイス42から読み出して、第1の記憶制御装置20に送信すると共に(ステップS33)、正常に上記データの読み出しが完了した旨を報告する(ステップS35)。第1の記憶制御装置20は、図8(b)に示すように、第2の記憶制御装置40から受信したデータを、キャッシュメモリ24の所定の場所に格納させる(ステップS34)。

## 【0085】

次に、第1の記憶制御装置20は、キャッシュメモリ24に格納されたデータを読み出し、アドレス変換を行った後、LUN103等を介してホスト装置10にデータを送信すると共に(ステップS36)、ホスト装置10に対して上記データの読み出し完了報告を行う(ステップS37)。これらデータ読み出し時の一連の処理動作では、図7と共に説明した変換操作が逆向きで行われる。

## 【0086】

図 8 では、ホスト装置 1 0 からの要求に応じて、第 2 の記憶制御装置 4 0 からデータを読み出し、キャッシュメモリ 2 4 に保存するかのように示している。しかし、これに限らず、外部の L D E V 5 0 に記憶されているデータの全部又は一部を、予めキャッシュメモリ 2 4 に記憶させておくこともできる。この場合、ホスト装置 1 0 からの読み出しコマンドに対し、直ちにキャッシュメモリ 2 4 からデータを読み出してホスト装置 1 0 に送信することができる。

#### 【 0 0 8 7 】

以上詳細に説明したとおり、本実施形態によれば、外部の記憶デバイス 4 2（正確には、外部の L D E V 5 0）を V D E V 1 0 1 にマッピングする構成のため、外部の論理ボリュームを内部の論理ボリュームのように、或いは、外部の記憶デバイス 4 2 を仮想的な内部記憶デバイスのように、取り扱うことができる。従って、第 2 の記憶制御装置 4 0 が、ホスト装置 1 0 と直接接続することができない旧型の装置である場合でも、新型の第 1 の記憶制御装置 2 0 が介在することにより、旧型装置の記憶資源を第 1 の記憶制御装置 2 0 の記憶資源として再利用し、ホスト装置 1 0 に提供することができる。これにより、旧型の記憶制御装置を新型の記憶制御装置 2 0 に統合して記憶資源を有効に利用できる。

#### 【 0 0 8 8 】

また、第 1 の記憶制御装置 2 0 が高性能・高機能の新型装置の場合、第 1 の記憶制御装置 2 0 の有する高性能なコンピュータ資源（キャッシュ容量や C P U 処理速度等）によって第 2 の記憶制御装置 4 0 の低性能を隠すことができ、外部の記憶デバイス 4 2 を活用した仮想的な内部ボリュームを用いて、高性能なサービスをホスト装置 1 0 に提供することができる。

#### 【 0 0 8 9 】

更に、外部の記憶デバイス 4 2 上に構築される L D E V 5 0 に、例えば、ストライピング、拡張、分割、RAID等の機能を追加して使用することができる。従って、外部のボリュームを直接 L U N 1 0 3 にマッピングする場合に比較して、利用の自由度が高まり、使い勝手が向上する。

#### 【 0 0 9 0 】

また、外部の論理ボリュームを内部の論理ボリュームのように使用できるため、通常の内部ボリュームである L D E V 1 0 2 に対して第 1 の記憶制御装置 2 0 が利用可能な諸機能を、仮想的な内部ボリューム（L D E V 5 0 に接続される L D E V）に対しても適用できる。利用可能な諸機能としては、例えば、M R C F、リモートコピー、C V S、L U S E等を挙げることができる。ここで、M R C F（Multiple RAID Coupling Feature）とは、ホスト装置 1 0 をデータが経由することなく（ホストフリー）、論理ボリュームのレプリカを作成可能な機能である。リモートコピーとは、ローカルサイトに設置されたプライマリボリュームとリモートサイトに設置されたセカンダリボリュームとの記憶内容を同期させる機能である。C V S（Customizable Volume Size）とは、論理ボリュームのサイズを標準的なサイズ以内の任意のサイズで設定可能な可変ボリューム機能である。L U S E（LU Size Expansion）とは、複数の論理ボリュームを 1 つの論理ボリュームに統合し、ホスト装置 1 0 から認識可能な L U N 数を少なくさせる L U N サイズ拡張機能である。

#### 【 0 0 9 1 】

更に、外部の論理ボリュームから構築される V D E V 1 0 1 を複数の L D E V 1 0 2 に夫々接続することができるため、ホスト装置 1 0 を各 L D E V 1 0 2 の L U N 1 0 3 に夫々接続することにより、交代パス構造を得ることができ、また、負荷分散効果も得ることができる。

#### 【 0 0 9 2 】

また、照会コマンドにより第 2 の記憶制御装置 4 0 の有する交代パス構造まで把握して V D E V 1 0 1 にマッピングするため、第 2 の記憶制御装置 4 0 の有する交代パス構造も第 1 の記憶制御装置 2 0 において継承することができるから、記憶システムの冗長性を高めることができる。

#### 【 0 0 9 3 】

これに対し、背景技術の欄で挙げた従来技術（特許文献2）は、ローカルな記憶デバイス、即ち、記憶制御装置が直接支配下に置いている記憶デバイスのボリュームをセクタ単位で再構成するだけのものであり、本実施形態のように、外部の記憶デバイスを仮想的な内部の記憶デバイスとして取り扱うものではない。また、他の従来技術（特許文献3）は、ローカルな記憶デバイスの記憶容量に合わせてエリアのサイズを最適化するもので、本実施形態のように、外部の記憶デバイス42を仮想的な内部記憶デバイスとして利用可能な技術ではない。

#### 【0094】

ところで、図1で示したホスト装置10のアプリケーションプログラム11が、第2の記憶制御装置40からのインクアイリ（照会）情報を取得しようとするに際して、図9（a）に示すように、ホスト装置10と第2の記憶制御装置40との間の接続が不良なため、ホスト装置10と第2の記憶制御装置40との間のデータの通信性能が悪い場合がある。また、上記インクアイリ情報の取得に際して、図9（b）に示すように、ホスト装置10と第2の記憶制御装置40との間を接続することができないため、アプリケーションプログラム11が第2の記憶制御装置40内のインクアイリ情報をサポートできない場合もある。

#### 【0095】

このような場合には、図9（c）に示すように、第2の記憶制御装置40内の（例えば、図1で示したような）記憶デバイス（42）を、第1の記憶制御装置20内の（例えば、図1で示したような）記憶デバイス（31）を仮想化したデバイスである仮想デバイス（VDEV101）にマッピングする。これにより、第1の記憶制御装置20は、第2の記憶制御装置40内の記憶デバイス（42）を、自身の仮想デバイス（VDEV101）としてホスト装置10に認識させることができるので、第1の記憶制御装置20から見て、外部デバイス資源（データ格納エリア）である第2の記憶制御装置40内の記憶デバイス（42）を、ホスト装置10と接続することの可能なデバイスの資源として利用することができる。アプリケーションプログラム11による第2の記憶制御装置40内のインクアイリ情報のサポートにおいても同様である。

#### 【0096】

図10は、本発明の第2の実施形態に係る第1の記憶制御装置10が第2の記憶制御装置40から受け取ったインクアイリ情報を、ホスト装置10のアプリケーションプログラム11に渡すに際しての第1の記憶制御装置20内における処理流れを示すブロック図である。

#### 【0097】

図10において、コマンドデバイス121とは、特殊なLU（ロジカルユニット）のことである。共有メモリ（SM）25には、コマンドデバイス121が特殊なLUであることを示すコマンドデバイス情報123、及びホスト装置10（のアプリケーションプログラム11）に転送するための、（外部ストレージである）第2の記憶制御装置40からのインクアイリ（照会）情報が記憶されている。

#### 【0098】

チャンネルプロセッサ（CHP）125は、図1で示したチャンネルアダプタ（CHA）21内に含まれるもので、ローカルメモリ（LM）127を備えており、共有メモリ（SM）25からコマンドデバイス情報123や、第2の記憶制御装置40からのインクアイリ情報等を読み出す。チャンネルプロセッサ125は、アプリケーションプログラム11から、コピー指示内容（第2の記憶制御装置40からのインクアイリ情報を取得するための処理内容）が、コマンドデバイス121に向けて発行されたSCSIのライト（書き込み）コマンドとして第1の記憶制御装置10へ転送されてくると、該ライトコマンドを受け付ける。そして、チャンネルプロセッサ125は、上記ライトコマンドのコマンドデバイス121への書き込みのため、上記コピー指示内容をローカルメモリ127に直接取り込む。

#### 【0099】

チャンネルプロセッサ125は、上記コピー指示内容を、ディスクアダプタ（DKA）を

通じてディスクへ書き込まずに解析を行い、この解析の結果に従ってオープンリモートコピー／MRCFの指示を、オープンリモートコピー／MRCF制御部129に出力する。チャンネルプロセッサ125は、オープンリモートコピー／MRCF制御部129から、(上記解析の結果に基づいた)オープンリモートコピー／MRCFの指示内容に問題が無いとの判断結果を受領した場合には、ホスト装置10に対するライトコマンドの終了ステータス報告として、ライト応答を正常とする旨の返信を行う。そして、続くホスト装置10によるインクアイリ情報のリード(読み込み)に対処するため、チャンネルプロセッサ125は、共有メモリ25からホスト装置10に返信すべきインクアイリ情報を検索して、第1の記憶制御装置20からのホスト装置10への出力データを作成する。

#### 【0100】

チャンネルプロセッサ125は、オープンリモートコピー／MRCF制御部129から、(上記解析の結果に基づいた)オープンリモートコピー／MRCFの指示内容に問題がある(矛盾がある)との判断結果を受領した場合には、ホスト装置10に対するライトコマンドの終了ステータス報告として、ライト応答をエラーとする旨の返信を行う。チャンネルプロセッサ125は、アプリケーションプログラム11から、上記インクアイリ情報に係わる上記出力データを読み込むため、コマンドデバイス121に向けて発行されたSCSIのリードコマンドが第1の記憶制御装置10へ転送されてくると、該リードコマンドを受け付ける。そして、チャンネルプロセッサ125は、ホスト装置10(のアプリケーションプログラム11)が先に上記ライトコマンドを発行したホスト装置10(のアプリケーションプログラム11)であるかどうかを確認できた場合には、上記出力データをホスト装置10に対するリード応答の戻り値として、ホスト装置10に返信する。

#### 【0101】

オープンリモートコピー／MRCF制御部129は、チャンネルプロセッサ125とは別の、独立したプロセッサであり、チャンネルプロセッサ125からのオープンリモートコピー／MRCFの指示内容が正しい要求であるかどうか(OK/NG)を判断して、その判断結果を、チャンネルプロセッサ125に通知する。

#### 【0102】

図11は、本発明の第2の実施形態に係る第1の記憶制御装置10が第2の記憶制御装置40から受け取ったインクアイリ情報を、ホスト装置10のアプリケーションプログラム11に渡すに際しての第1の記憶制御装置20内における処理流れを示すフローチャートである。

#### 【0103】

図11において、まず、ホスト装置10が、第1の記憶制御装置20のコマンドデバイス121に向けてインクアイリ情報等を取得するための指示内容(コピー指示内容)を(SCSIの)ライトコマンドとして発行すると(ステップS151)、チャンネルプロセッサ125は、該コマンドを受け付ける。そして、上記ライトコマンドのコマンドデバイス121への書き込みのため、上記コピー指示内容をローカルメモリ127に直接取り込む(ステップS152)。次に、チャンネルプロセッサ125は、上記コピー指示内容の解析を行い、この解析の結果に従ってオープンリモートコピー／MRCFの指示を、オープンリモートコピー／MRCF制御部129に出力する(ステップS153)。

#### 【0104】

オープンリモートコピー／MRCF制御部129から、(上記解析の結果に基づいた)オープンリモートコピー／MRCFの指示内容に問題が無いとの判断結果を受領した場合には、チャンネルプロセッサ125は、ホスト装置10に対するライトコマンドの終了ステータス報告として、ライト応答を正常とする旨の返信を行う(ステップS154)。それに引き続いて、ホスト装置10によるインクアイリ情報のリードに対処するため、チャンネルプロセッサ125は、共有メモリ25からホスト装置10に返信すべきインクアイリ情報を検索し、第1の記憶制御装置20からのホスト装置10への出力データを作成する(ステップS155)。

#### 【0105】

一方、チャネルプロセッサ125は、オープンリモートコピー/MRCF制御部129から、上記指示内容に問題がある（矛盾がある）との判断結果を受領した場合には、ホスト装置10に対するライトコマンドの終了ステータス報告として、ライト応答をエラーとする旨の返信を行う（ステップS156）。

#### 【0106】

次に、アプリケーションプログラム11から、上記インクアイリ情報に係わる上記出力データを読み込むため、コマンドデバイス121に向けて発行されたSCSIのリードコマンドが第1の記憶制御装置10へ転送されてくると（ステップS157）、該リードコマンドを受け付ける（ステップS158）。そして、チャネルプロセッサ125は、ホスト装置10（のアプリケーションプログラム11）が先に上記ライトコマンドを発行したホスト装置10（のアプリケーションプログラム11）であるかどうかを確認できた場合には、上記出力データをホスト装置10に対するリード応答の戻り値として、ホスト装置10に返信する（ステップS159）。

#### 【0107】

ところで、上述したホスト装置10、及び第1の記憶制御装置20から見て外部のホスト装置（以下、「外部ホスト装置」と表記する）（図1等では、図示省略）（のアプリケーションプログラム）が、同じく外部の記憶制御装置（以下、「外部記憶制御装置」という）（図1等で示した第2の記憶制御装置40）からのインクアイリ情報を取得しようとする際に、外部記憶制御装置の構成が、同じく外部ホスト装置からの（リモートコピーの）コマンドに対応不能な場合がある。このような場合には、図12（a）において符号161で示すような、外部ホスト装置163からの（リモートコピーの）コマンドに対応可能な構成の外部記憶制御装置を新たに設ける。これにより、外部ホスト装置163のアプリケーションプログラム165は、上記対応不能な外部記憶制御装置167からのインクアイリ情報を、外部記憶制御装置161を通じて取得することが可能になる。

#### 【0108】

図12（a）に示すように、外部記憶制御装置161には、第1の記憶制御装置20内の（例えば、図1で示したような）記憶デバイス（31）を、外部記憶制御装置161内の記憶デバイス（例えば、図1で示したような第2の記憶制御装置40内の記憶デバイス42）を仮想化したデバイスである仮想デバイス（VDEV）162にマッピングする。これにより、外部記憶制御装置161は、第1の記憶制御装置20内の記憶デバイス（31）を、自身の仮想デバイス（VDEV）162として外部ホスト装置163に認識させることができる。

#### 【0109】

図12（b）に示すように、外部ホスト装置163と外部記憶制御装置167とが接続不能であるが故に、アプリケーションプログラム165による外部記憶制御装置167内のインクアイリ情報のサポートを直接行うことはできない。そのため、図12（b）に示すように、アプリケーションプログラム165による上記インクアイリ情報のサポートを可能ならしめるべく、外部記憶制御装置167内の記憶デバイスを、外部記憶制御装置161内の記憶デバイスを仮想化したデバイスである仮想デバイス（VDEV）162にマッピングすることになる。（なお、この仮想デバイス162には、外部記憶制御装置161内の記憶デバイス（のデータ）もコピーされる。）しかし、外部記憶制御装置161内においてさえ、該装置161内コピーにおけるコピー先の記憶デバイスが不明であるために、上記マッピングが、外部記憶制御装置161内の何れの記憶デバイスに対して行われたか不明であるから、このような状態では、アプリケーションプログラム165側から外部記憶制御装置161に対して、インクアイリ情報等のコピーの指示を発行するのは不可能である。

#### 【0110】

そこで、本発明の第3の実施形態では、図13（a）に示すように、外部記憶制御装置161が、該記憶制御装置161内の、上記マッピングが施された記憶デバイス162に係わるデバイスインクアイリ情報から得られるIDを、該記憶制御装置161内に蓄積し

ておく。より具体的に説明すると、外部記憶制御装置 161 は、図 13 (b) に示すように、該記憶制御装置 161 内の仮想デバイス (VDEV) 162 と、該仮想デバイス (VDEV) 162 にマッピングした外部記憶制御装置 167 内の記憶デバイス 168 との対応関係を示すマッピングテーブルを、図 1 において符号 25 で示したのと同様な共有メモリに格納している。このマッピングテーブルには、外部記憶制御装置 167 内の上記記憶デバイス 168 に係わる情報である、デバイスインクアイリ情報 (図 13 (a) では、「外部デバイス情報」と表記している) が格納されており、デバイスインクアイリ情報には、デバイス識別情報 (ID) が含まれている。デバイス識別情報 (ID) は、デバイスインクアイリ情報中から、外部記憶制御装置 167 内の上記記憶デバイス 168 に係わるユニークな部分を抜粋して生成したものである。なお、デバイス識別情報 (ID) としては、例えば DF では、Vender ID, Product ID, Vender Specific 内のシリアル番号、LDEV 番号等が用いられる。

#### 【0111】

図 13 (c) は、標準インクアイリデータの抜粋を示すもので、図 13 (c) で示した 3 つのデータがデバイス識別情報 (ID) を生成するために用いるユニークな値である。

#### 【0112】

再び、図 13 (a) に戻って、外部記憶制御装置 161 と外部記憶制御装置 167 との間においては、ライト、リード、インクアイリ等の SCSI コマンドの授受が可能である。上述したデバイス情報 (ID) は、外部記憶制御装置 161 内のチャンネルプロセッサ (CHP) により、予め上記デバイスインクアイリ情報中から抽出されて (共有メモリに格納されている) マッピングテーブルに蓄積される。

#### 【0113】

(外部ホスト装置 163 内の) アプリケーションプログラム 165 は、例えば外部記憶制御装置 161 のチャンネルプロセッサ (CHP) を介して外部記憶制御装置 167 から上記デバイスインクアイリ情報を取得し、該インクアイリ情報から上述したデバイス識別情報 (ID) を生成する。そして、アプリケーションプログラム 165 は、このデバイス識別情報 (ID) と、コピー元である外部記憶制御装置 161 内の上述した特定の記憶デバイス 164 に係わるデバイス識別情報 (ID) とをコピー先のキーとして、外部記憶制御装置 161 に対し、インクアイリ情報等のコピーの指示を発行する。

#### 【0114】

外部記憶制御装置 161 は、アプリケーションプログラム 165 からのコピーの指示を受け付けると、例えば上述したチャンネルプロセッサ (CHP) が、上記 ID に基づいて、外部記憶制御装置 161 内の上述した仮想デバイス (VDEV) 162 として割り当てられた記憶デバイスを検索する。そして、検索された記憶デバイス 162、及び該記憶制御装置 161 内の他の記憶デバイスからのコピーが実行され、検索された記憶デバイス 162 のインクアイリ情報、及び該記憶制御装置 161 内の他の記憶デバイスのインクアイリ情報が、アプリケーションプログラム 165 によって取得される。

#### 【0115】

これによって、アプリケーションプログラム 165 は、外部記憶制御装置 167 内の記憶デバイスが外部記憶制御装置 161 内の記憶デバイスに対して、どのような仮想的な割り当てがなされているかを知ること無く、外部記憶制御装置 161 内の目的とする記憶デバイスからのコピーを実行することができ、所望のインクアイリ情報を取得することができる。

#### 【0116】

図 14 は、本発明の第 3 の実施形態に係る外部ホスト装置 163 と外部記憶制御装置 161 との間にインクアイリ情報の要求・応答をする場合の、外部ホスト装置 163 内、及び外部記憶制御装置 161 内の処理流れを示すフローチャートである。

#### 【0117】

図 14 において、まず、外部ホスト装置 163 が、外部記憶制御装置 161 のチャンネルアダプタ (図示しない) を介して、外部記憶制御装置 161 にログインする (ステップ S

171)。この処理動作の後に、外部記憶制御装置161が、外部ホスト装置163のログインに対して応答を返すことにより、ログインが完了する(ステップS172)。次に、外部ホスト装置163は、例えば、SCSI規格で定められているインクアイリコマンドを外部記憶制御装置161に送信して、外部記憶制御装置161が持つ記憶デバイスの詳細について応答を求める(ステップS173)。

#### 【0118】

ここで、インクアイリコマンドとは、照会先の装置の種類及び構成を明らかにするために用いられるもので、照会先装置の有する(記憶)階層を透過してその物理的構造を把握することができる。インクアイリコマンドを使用することにより、外部ホスト装置163は、例えば、装置名、デバイスタイプ、製造番号(プロダクトID)、LDEV番号、各種バージョン情報、ベンダID等の、交代パスを意識できるデバイス情報を、外部記憶制御装置161から取得することができる(ステップS174)。外部記憶制御装置161は、外部ホスト装置163から問合せされた情報を、外部ホスト装置163に送信し、応答する(ステップS175)。この処理が完了することで、一連の処理動作が終了することになる。

#### 【0119】

図15は、本発明の第4の実施形態に係る外部ホスト装置163が外部記憶制御装置161から受け取ったインクアイリ情報を、ホスト装置10(内部ホスト装置)のアプリケーションプログラム(11)に渡すときの、外部ホスト装置163内、及び外部記憶制御装置161内の処理流れを示すフローチャートである。なお、図15に示す処理流れは、図10で記載した各機能ブロックの処理動作を表しており、図12、及び図13で記載したシステム構成を前提としている。

#### 【0120】

図15において、まず、外部ホスト装置163が、外部記憶制御装置161のコマンドデバイス(図示しない)に向けてインクアイリ情報等を登録するための指示内容(コピー指示内容)を(SCSIの)ライトコマンドとして発行すると(ステップS181)、外部記憶制御装置161のチャンネルプロセッサ(CHP)(図示しない)は、該コマンドを受け付ける。そして、上記ライトコマンドのコマンドデバイス(図示しない)への書き込みのため、上記コピー指示内容をローカルメモリ(図示しない)に直接取り込む(ステップS182)。次に、チャンネルプロセッサ(図示しない)は、上記コピー指示内容の解析を行い、この解析の結果、上記コピー指示内容に問題が無いと判断した場合には、チャンネルプロセッサ(図示しない)は、外部ホスト装置163に対するライトコマンドの終了ステータス報告として、ライト応答を正常とする旨の返信を行う(ステップS183)。それに引き続いて、外部記憶制御装置161では、上記インクアイリ情報等を共有メモリ(SM)(図示しない)に保持する処理を行う(ステップS184)。この処理が完了することで、一連の処理動作が終了することになる。

#### 【0121】

図16は、本発明の第5の実施形態に係る第1の記憶制御装置20(図1で示した)が、ホスト装置10内のアプリケーションプログラム11(図1で示した)からの指示に応じて、第1の記憶制御装置20内の記憶領域と外部記憶制御装置である第2の記憶制御装置40(図1で示した)内の記憶領域との間でペアを形成するに際しての処理流れを示すフローチャートである。なお、図16に示す処理流れは、図10で記載した各機能ブロックの処理動作を表している。

#### 【0122】

図16において、まず、ホスト装置10が、第1の記憶制御装置20のコマンドデバイス(121)(図10で示した)に向けてリモートコピーペア形成指示のための情報(リモートコピーペア形成指示内容)を(SCSIの)ライトコマンドとして発行すると(ステップS191)、第1の記憶制御装置20では、このライトコマンドを受領する。チャンネルアダプタ(CHA)(21)(図1で示した)は、上記ライトコマンドのコマンドデバイス(121)への書き込みのため、上記リモートコピーペア形成指示内容をローカル



メモリ (127) (図10で示した) に直接取り込む (ステップS192)。次に、チャネルアダプタ (21) は、上記リモートコピーペア形成指示内容の解析を行い、この解析の結果、上記リモートコピーペア形成指示内容に問題が無いと判断した場合には、ディスクアダプタ (DKA) (22) (図1で示した) に対してリモートコピーペア形成指示を行う (ステップS193)。

#### 【0123】

このリモートコピーペア形成指示を受け付けると、ディスクアダプタ (22) は、該指示に従って上記ペアの作成、及び初期データ (インクアイリデータ) のコピーを開始する (ステップS194)。そして、ディスクアダプタ (22) は、上記ペアの作成の過程において問題が生じなかったと判断した場合には、ペア作成が正常に終了した旨の応答をチャネルアダプタ (21) に対して返信し (ステップS195)、問題が生じたと判断した場合には、ペア作成にエラーが生じた旨の応答をチャネルアダプタ (21) に対して返信する (ステップS196)。

#### 【0124】

一方、チャネルアダプタ (21) は、ペア作成が正常に終了した旨の応答がディスクアダプタ (22) から返信されてきた場合には、ホスト装置 10 に対するリモートコピーペア形成指示コマンドの終了ステータス報告として、リモートコピーペアの形成が正常に終了した旨の応答を返信する (ステップS197)。しかし、ホスト装置 10 からの上記指示内容に矛盾があったり、ディスクアダプタ (22) から上記エラーの報告が返信されてきた場合には、チャネルアダプタ (21) は、リモートコピーの形成においてエラーが生じた旨の応答をホスト装置 10 に対して返信する (ステップS198)。

#### 【0125】

ところで、図17(a)に示すように、ホスト装置 10 (のアプリケーションプログラム 11) から第1の記憶制御装置 20 に発行されたコピー指示のコマンドに従って、第1の記憶制御装置 20 が、リモートコピー非対応の外部記憶制御装置 201 の記憶デバイス 209 から自装置 20 内の仮想デバイス 207 へのリモートコピーを実行する場合がある。また、このリモートコピーと共に、第1の記憶制御装置 20 が、自装置 20 内の記憶デバイス 205 から仮想デバイス 207 へのコピーを実行する場合もある。このようなコピーを第1の記憶制御装置 20 が実行している際に、外部記憶制御装置 201 が、直接接続されている外部ホスト装置 203 からアクセスされると、外部記憶制御装置 201 は、外部ホスト装置 203 からのアクセスを排除することができないため、該アクセスによって、上記コピーの内容が破壊される不具合が生じる。

#### 【0126】

そこで、本発明の第6の実施形態では、図17(b)に示すように、第1の記憶制御装置 20 がホスト装置 10 (のアプリケーションプログラム 11) から発行されるコピー指示のコマンドに従って、外部記憶制御装置 201 からのリモートコピーを開始する直前に、第1の記憶制御装置 20 から外部記憶制御装置 201 に対して reserve コマンドを発行することにより、上記不具合を未然に防止することとした。

#### 【0127】

図17(b)を用いて本発明の第6の実施形態をより具体的に説明する。第1の記憶制御装置 20 内のチャネルプロセッサ (CHP) (125) は、(ホスト装置 10 の) アプリケーションプログラム 11 からのコピー指示のコマンドを受け付けると、上記コピー、及びリモートコピーを開始する直前に、上記コマンド中の ID によって指示された外部記憶制御装置 201 内の記憶デバイス 209 に対して SCSI の reserve コマンドを発行する。この reserve コマンドを受け付けることにより、記憶デバイス 209 は、外部ホスト装置 203 からのアクセスを SCSI コマンドレベルで遮断することが可能になる。

#### 【0128】

記憶デバイス 205 からの仮想デバイス 207 に対するコピー、及び (外部記憶制御装置 201 の) 記憶デバイス 209 からの仮想デバイス 207 に対するリモートコピーが終了したことにより、チャネルプロセッサ (CHP) (125) が上記 reserve コマンドを



解除すれば、外部ホスト装置 203 からの外部記憶制御装置 201 に対するアクセスが再び可能になる。

#### 【0129】

以上説明したように、本実施形態によれば、(ホスト装置 10 の) アプリケーションプログラム 11 が、第 1 の記憶制御装置 20 に対してコピー指示のコマンドを発行するだけで、第 1 の記憶制御装置 20 を通じて外部記憶制御装置 201 に対する外部ホスト装置 203 からのアクセスを遮断することができ、それにより、コピーが破壊されるのを防止することができる。

#### 【0130】

ところで、図 18 (a) に示すように、記憶制御装置 221 が、ホスト装置 10 からのリモートコピー指示のコマンドを受け付ける構成になっていない場合や、外部記憶制御装置 225 が、ホスト装置 10 と通信可能な外部ホスト装置 223 からのリモートコピー指示のコマンドを受け付ける構成になっていない場合がある。これらの場合には、ホスト装置 10、及び外部ホスト装置 223 の何れにおいても、記憶制御装置 221 から外部記憶制御装置 225 へのリモートコピーを利用することができない。

#### 【0131】

そこで、本発明の第 7 の実施形態では、図 18 (b) に示すように、ホスト装置 10 側に、図 1 等で説明した第 1 の記憶制御装置 20 と同様の構成の記憶制御装置 227 を、内部記憶制御装置として設け、外部ホスト装置 223 側に、図 1 等で説明した第 2 の記憶制御装置 40 と同様の構成の記憶制御装置 229 を、外部記憶制御装置として設ける。そして、内部記憶制御装置 227 と外部記憶制御装置 229 とを介して、あたかも (内部) 記憶制御装置 221 (内のデータ) が外部記憶制御装置 225 側へリモートコピーされたように外部ホスト装置 223 に認識させるようにした。

#### 【0132】

図 18 (b) を用いて本発明の第 7 の実施形態をより具体的に説明する。内部記憶制御装置 227 において、複数の記憶デバイスの中から仮想化した仮想デバイス 231 に、(リモートコピー非対応の) 内部記憶制御装置 221 (のデータ) をコピーする。一方、外部記憶制御装置 229 においても、複数の記憶デバイスの中から仮想化した仮想デバイス 233 に、(リモートコピー非対応の) 外部記憶制御装置 225 (のデータ) をコピーする。

#### 【0133】

このようにして、外部記憶制御装置 225 (のデータ) がコピーされた後の外部記憶制御装置 229 の仮想デバイス 233 に対して、内部記憶制御装置 227 (に格納された内部記憶制御装置 221 からのデータ) が、リモートコピーされる。

#### 【0134】

上述した外部記憶制御装置 229 の仮想デバイス 233 は、外部ホスト装置 223 (のアプリケーションプログラム 235) から見ると、外部記憶制御装置 229 の記憶デバイスとしてしか認識されないため、外部ホスト装置 223 は、上記仮想デバイス 233 にコピーされたデータを利用することが可能になる。

#### 【0135】

以上説明したように、本実施形態によれば、リモートコピーをサポートしていない記憶制御装置 221、225 のリモートコピーが可能になり、且つ、リモートコピー制御をサポートするための (ホスト装置 10 の) アプリケーションプログラム 11 からのリモートコピー制御も可能になる。

#### 【0136】

以上、本発明の好適な実施形態を説明したが、これらは本発明の説明のための例示であって、本発明の範囲をこれらの実施形態にのみ限定する趣旨ではない。本発明は、他の種々の形態でも実施することが可能である。

#### 【図面の簡単な説明】

#### 【0137】

- 【図 1】 本発明の第 1 の実施例に係る記憶システムの全体構成を示すブロック図。
- 【図 2】 図 1 に記載の記憶システムの論理的構成の概要を示す模式図。
- 【図 3】 図 1 に記載の記憶システムの他の論理的構成の概要を示す模式図。
- 【図 4】 図 1 に記載の記憶システムのマッピングテーブルの概要を示す説明図。
- 【図 5】 図 1 に記載の記憶システムのマッピングテーブルを構築するための処理の流れを示す説明図。
- 【図 6】 図 1 に記載の記憶システムの内部ボリュームとして仮想化される外部の記憶デバイスにデータを書き込む場合の概念図。
- 【図 7】 図 1 に記載の記憶システムの書き込みデータのアドレス変換の様子を模式的に示す説明図。
- 【図 8】 図 1 に記載の記憶システムの内部ボリュームとして仮想化される外部の記憶デバイスからデータを読み出す場合の概念図。
- 【図 9】 本発明の第 2 の実施形態を実現するための記憶制御装置システムのシステム構成を示すブロック図。
- 【図 10】 本発明の第 2 の実施形態に係る第 1 の記憶制御装置が第 2 の記憶制御装置から受け取ったインクアイリ情報を、アプリケーションプログラムに渡すに際しての第 1 の記憶制御装置内における処理流れを示すブロック図。
- 【図 11】 本発明の第 2 の実施形態に係る第 1 の記憶制御装置が第 2 の記憶制御装置から受け取ったインクアイリ情報を、アプリケーションプログラムに渡すに際しての第 1 の記憶制御装置内における処理流れを示すフローチャート。
- 【図 12】 本発明の第 3 の実施形態を実現するための記憶制御装置システムのシステム構成を示すブロック図。
- 【図 13】 本発明の第 3 の実施形態を実現するための記憶制御装置システムのシステム構成を示すブロック図。
- 【図 14】 本発明の第 3 の実施形態に係る外部ホスト装置と外部記憶制御装置との間でインクアイリ情報の要求・応答をする場合の、外部ホスト装置内、及び外部記憶制御装置内の処理流れを示すフローチャート。
- 【図 15】 本発明の第 4 の実施形態に係る外部ホスト装置が外部記憶制御装置から受け取ったインクアイリ情報を、内部ホスト装置のアプリケーションプログラムに渡すときの、外部ホスト装置内、及び外部記憶制御装置内の処理流れを示すフローチャート。
- 【図 16】 本発明の第 5 の実施形態に係る第 1 の記憶制御装置が、内部ホスト装置のアプリケーションプログラムからの指示に応じて、第 1 の記憶制御装置内の記憶領域と外部記憶制御装置内の記憶領域との間でペアを形成するに際しての処理流れを示すフローチャート。
- 【図 17】 本発明の第 6 の実施形態を実現するための記憶制御装置システムのシステム構成を示すブロック図。
- 【図 18】 本発明の第 7 の実施形態を実現するための記憶制御装置システムのシステム構成を示すブロック図。

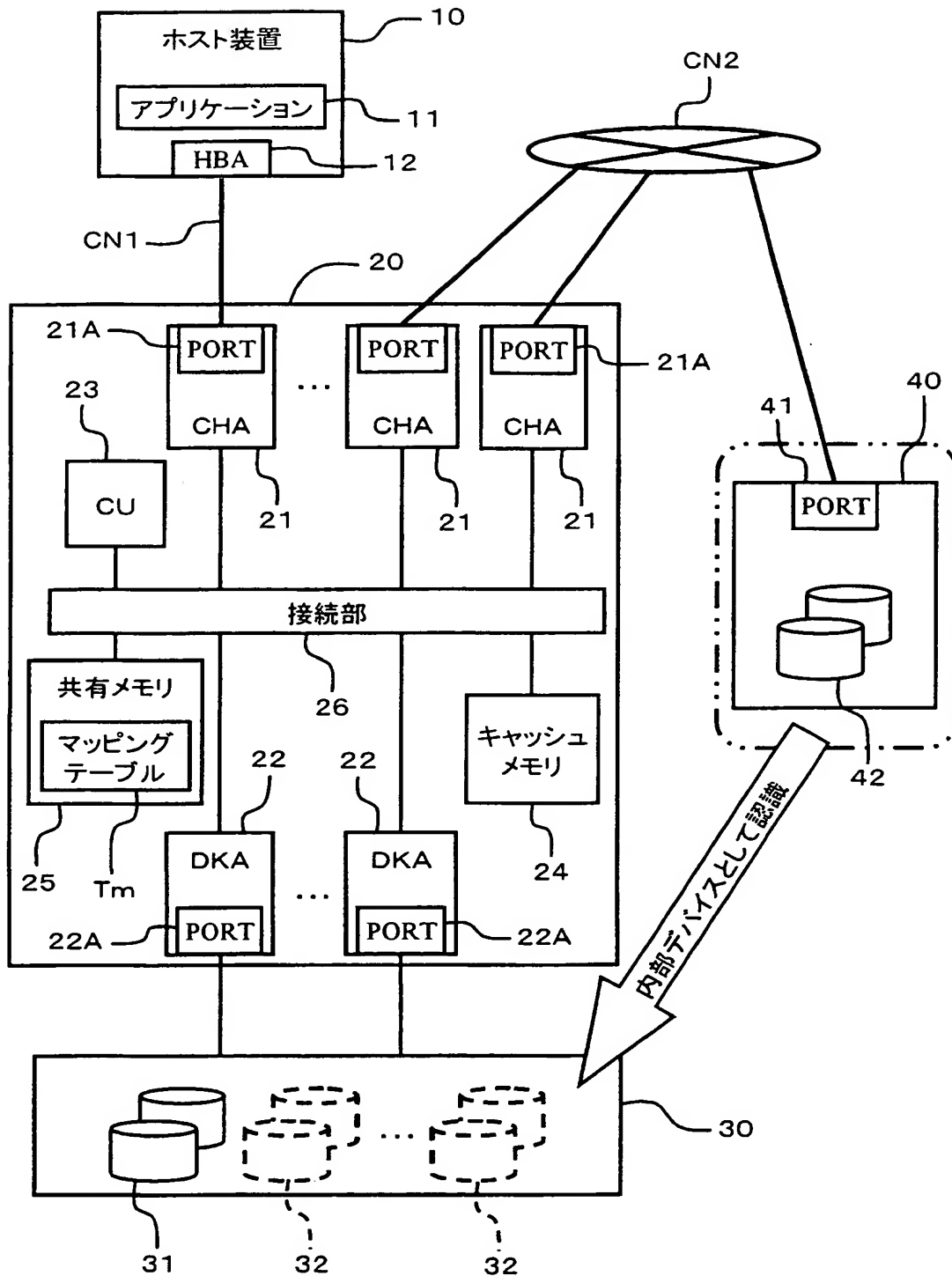
【符号の説明】

【0138】

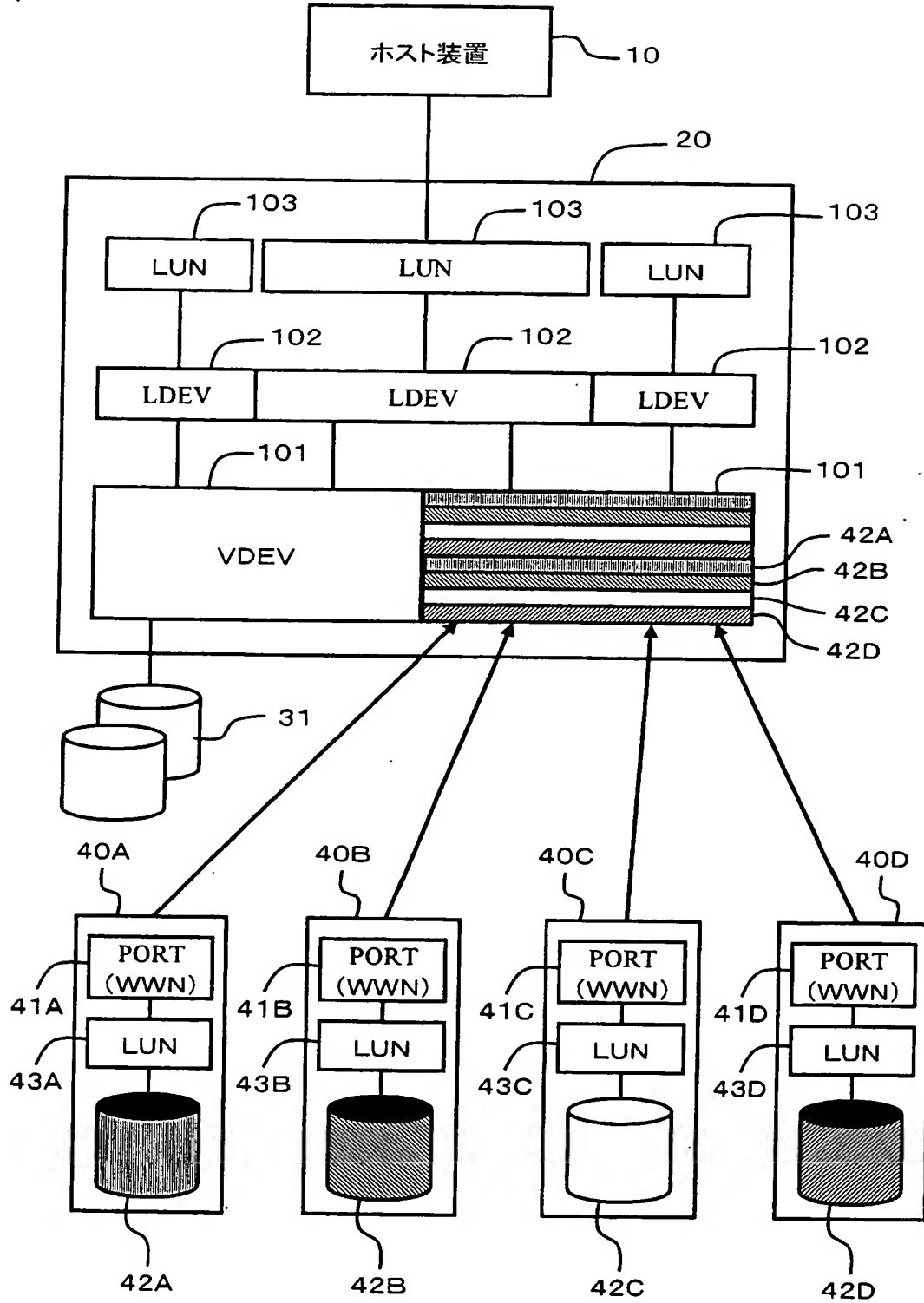
- 10 ホスト装置
- 11 アプリケーションプログラム
- 20 第 1 の記憶制御装置
- 21 チャンネルアダプタ
- 22 ディスクアダプタ
- 24 キャッシュメモリ
- 25 共有メモリ
- 31 記憶デバイス
- 32 記憶デバイス（仮想化された内部記憶デバイス）

- 4 0 第 2 の記憶制御装置
- 4 2 各記憶デバイス
- 1 2 1 コマンドデバイス
- 1 2 3 コマンドデバイス情報 1 2 3
- 1 2 5 チャンネルプロセッサ (CHP)
- 1 2 7 ローカルメモリ (LM)
- 1 2 9 オープンリモートコピー／MRCF 制御部

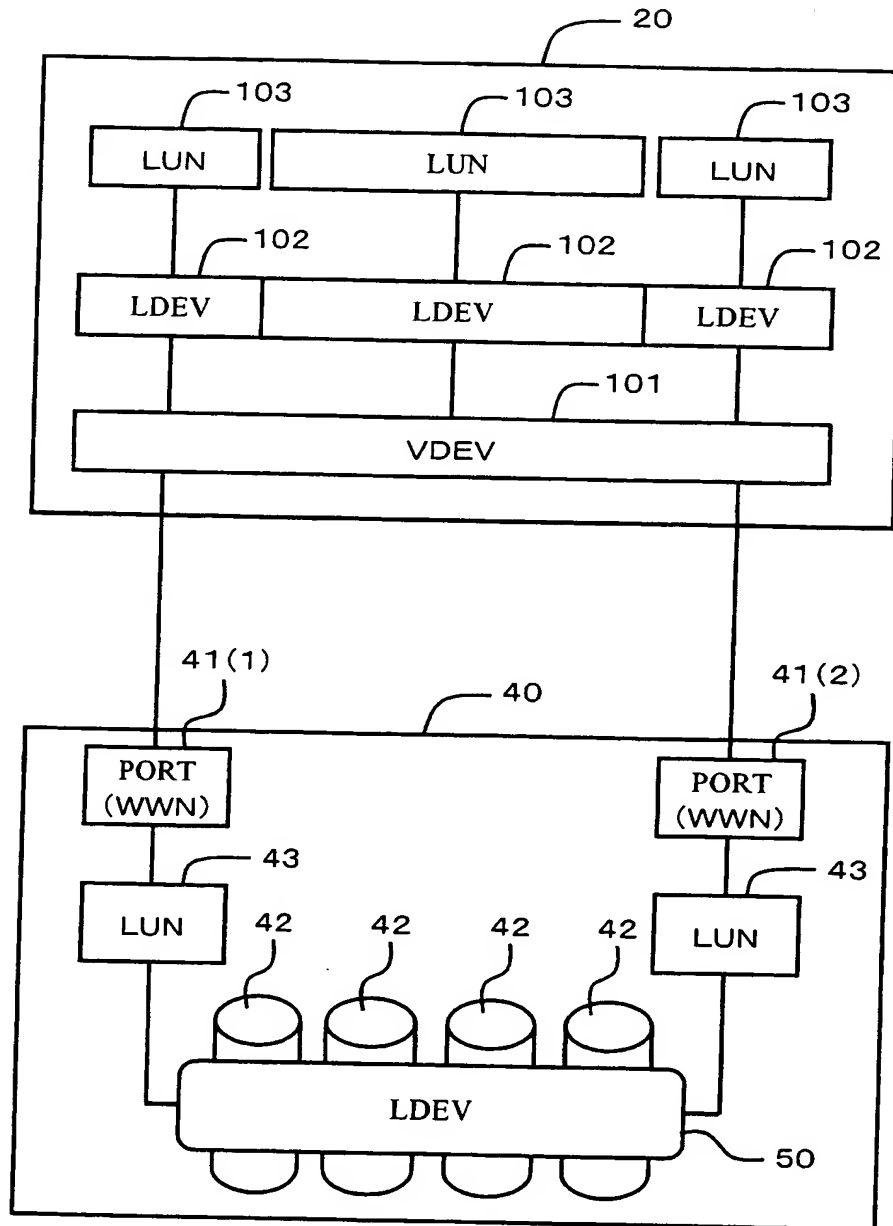
【書類名】 図面  
【図 1】



【図 2】



【図 3】

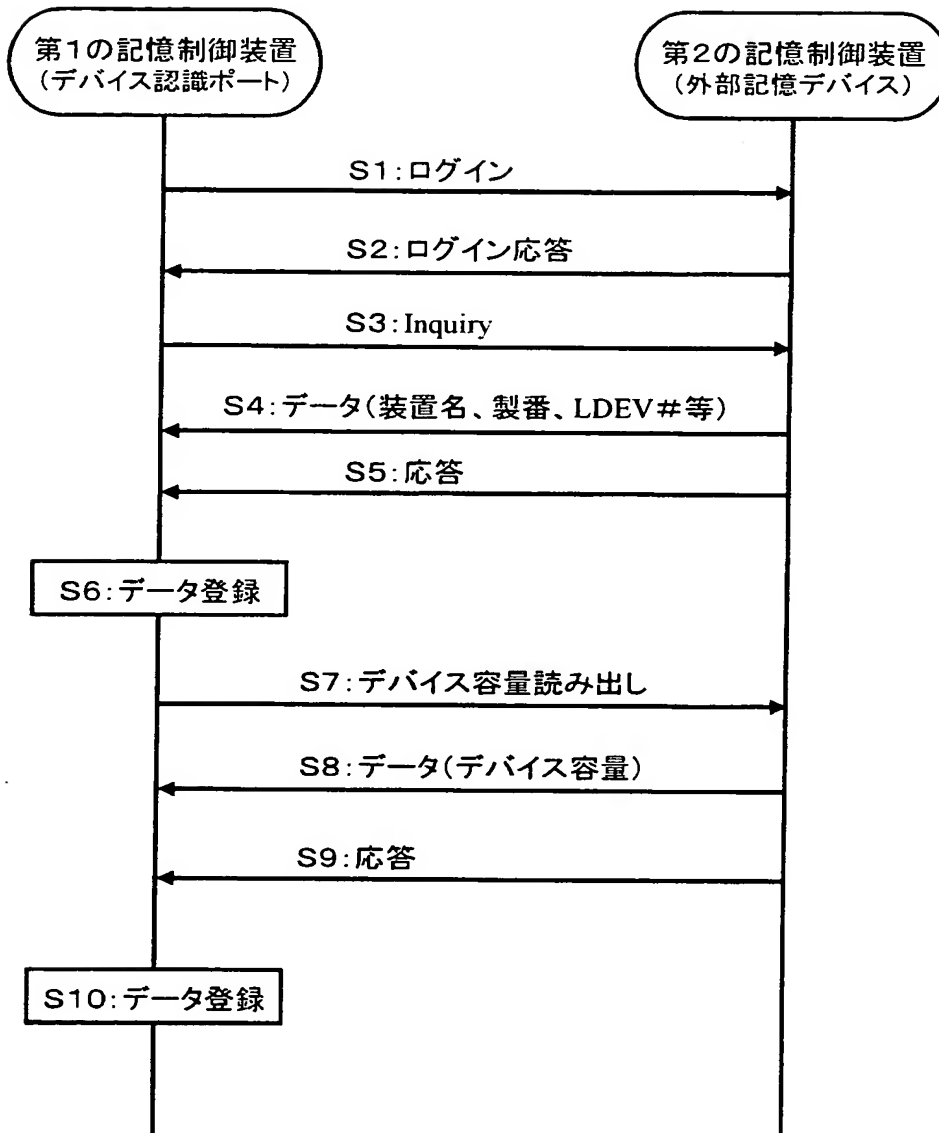


【図 4】

VDEV	外部デバイス情報				
	デバイス 識別情報	容量 (KB)	デバイス 種別	バス情報	
				WWN	LUN
0	DRFGTFNEIEK	657,456	DISK	0xAABBCCDD	0
1	ADRFGTNEIE	89,854	DISK	0xAABBEFF	3
2	GRRFFDDERT	—	TAPE	0x445566AAB	5
3	AABCCDDEE	5,544,223	DISK	0x77DE12345	6
				0x77DE12345	3
				0x377DE7890	5

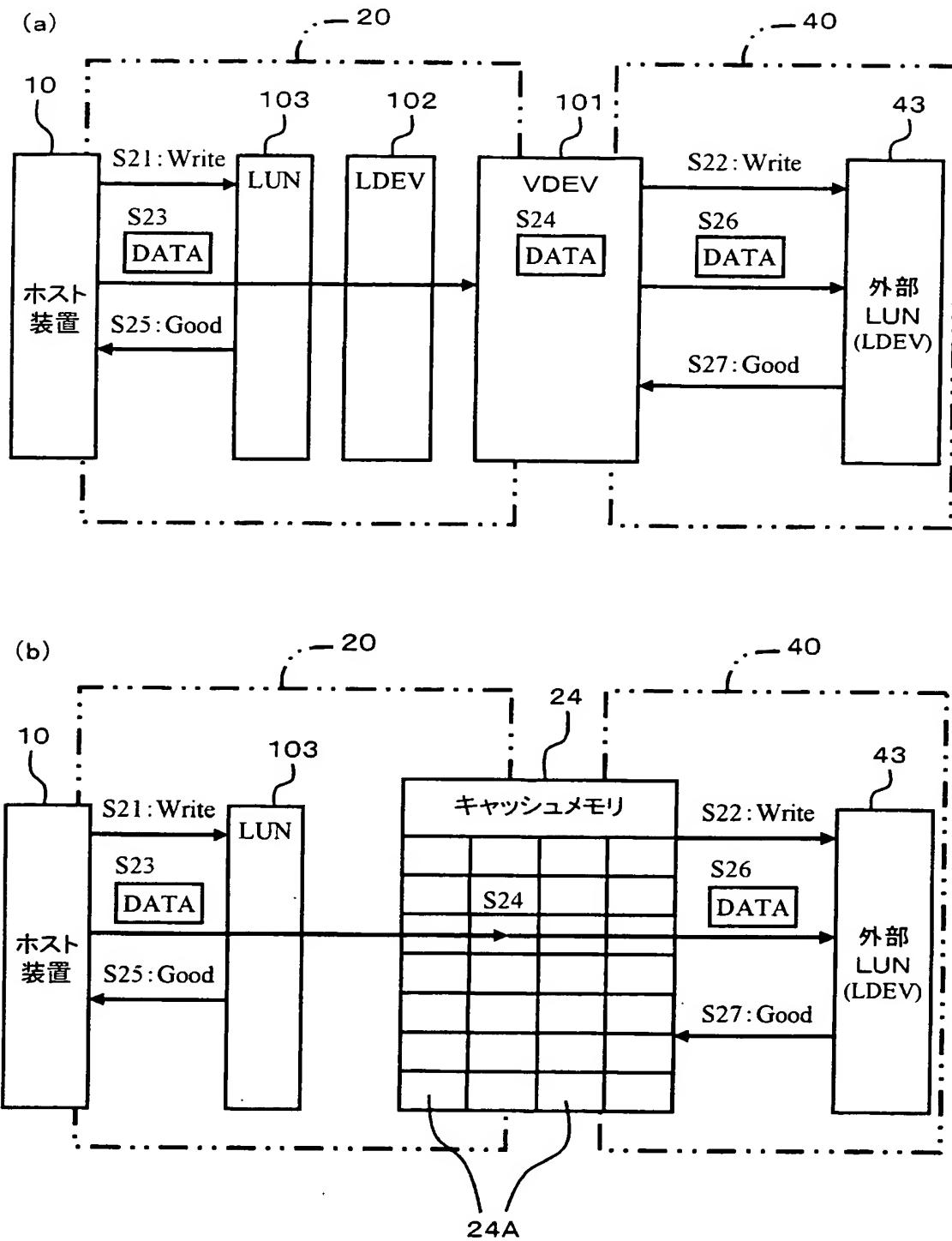
交代バス

【図5】

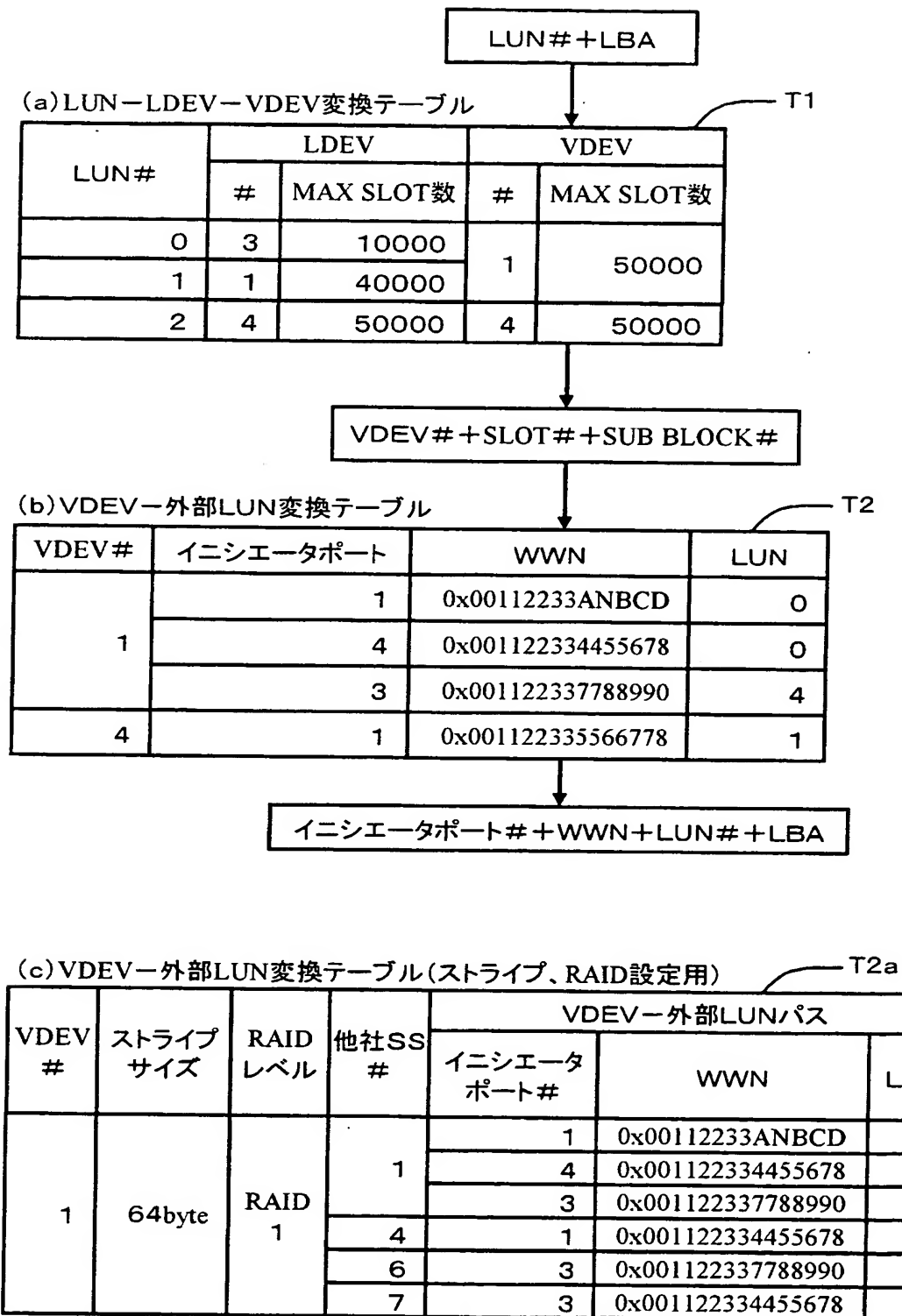




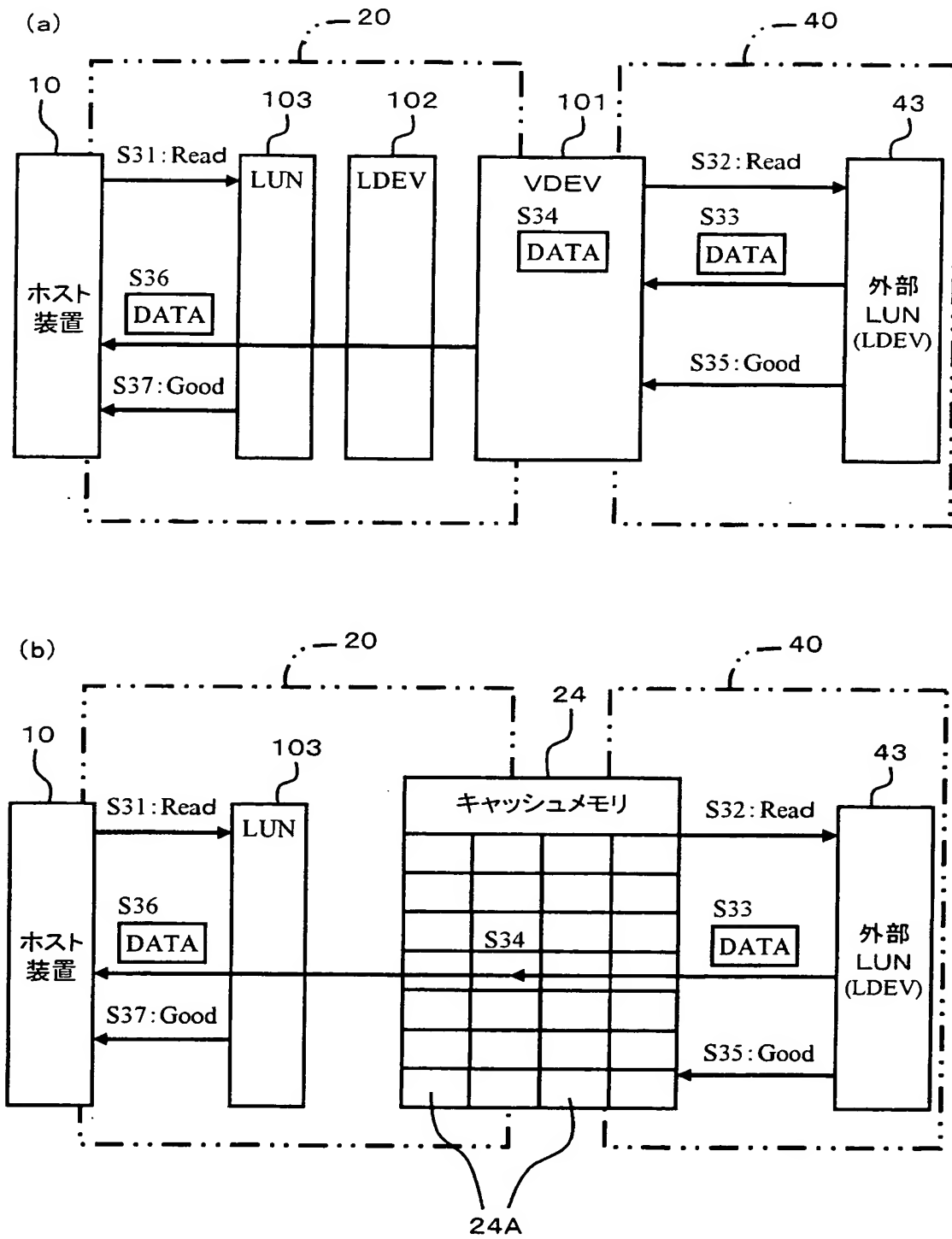
【図 6】



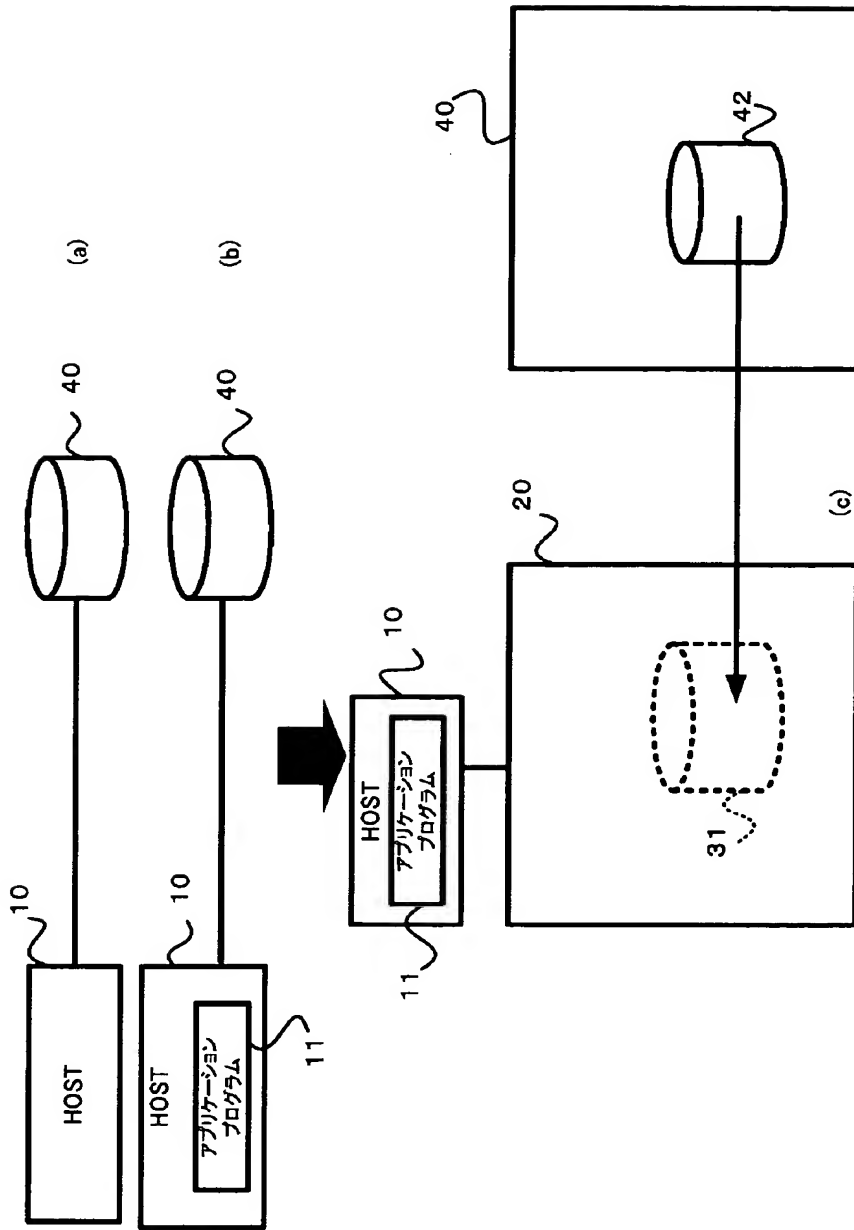
【図7】



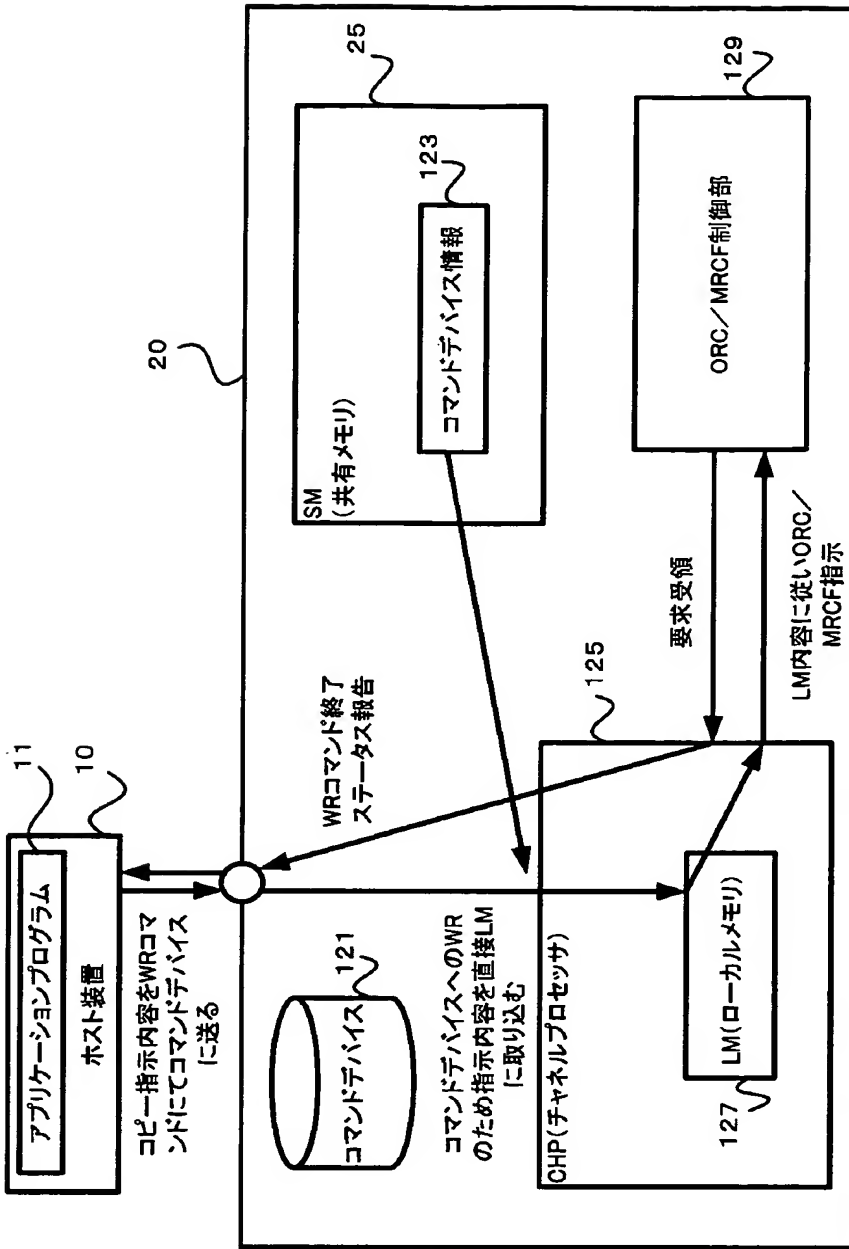
【図8】



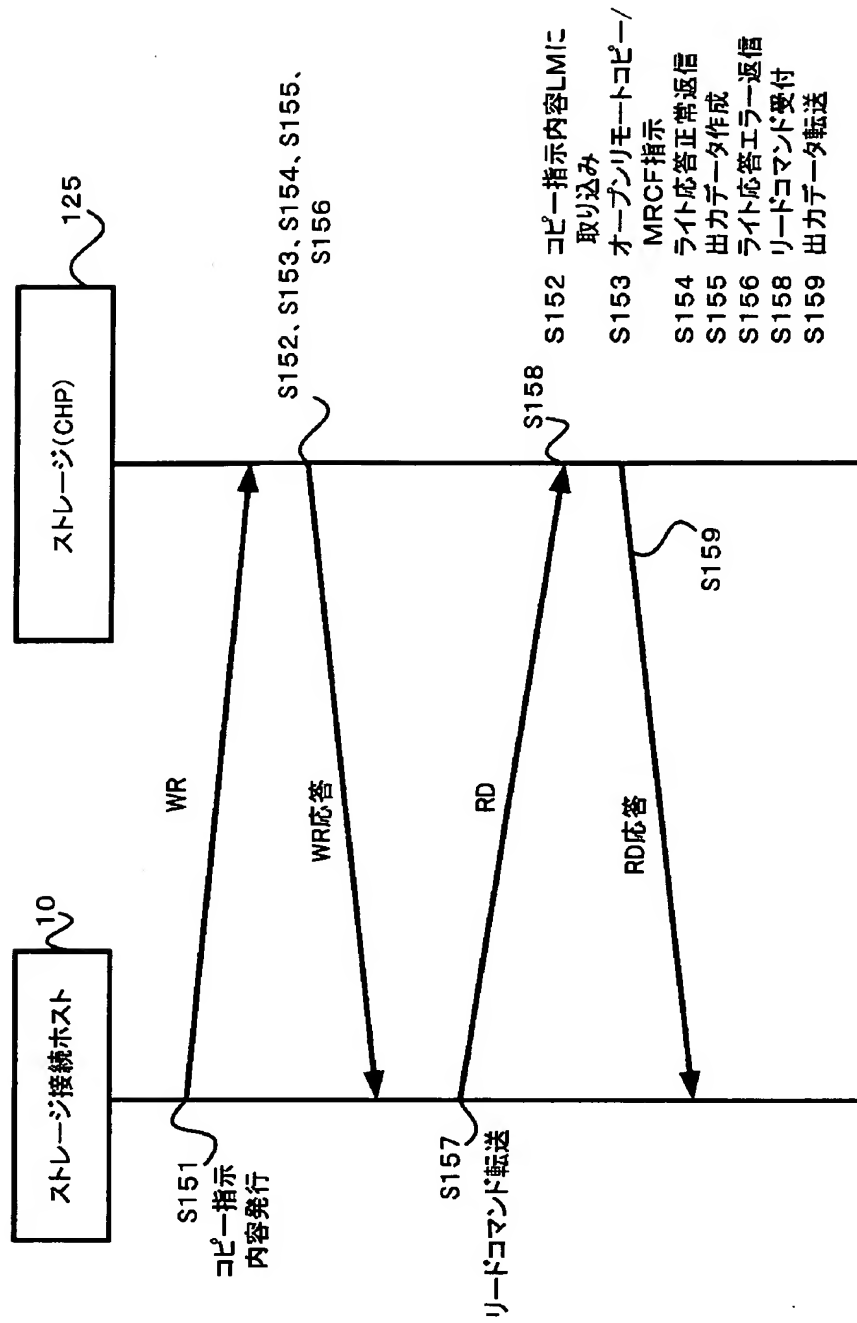
【図9】



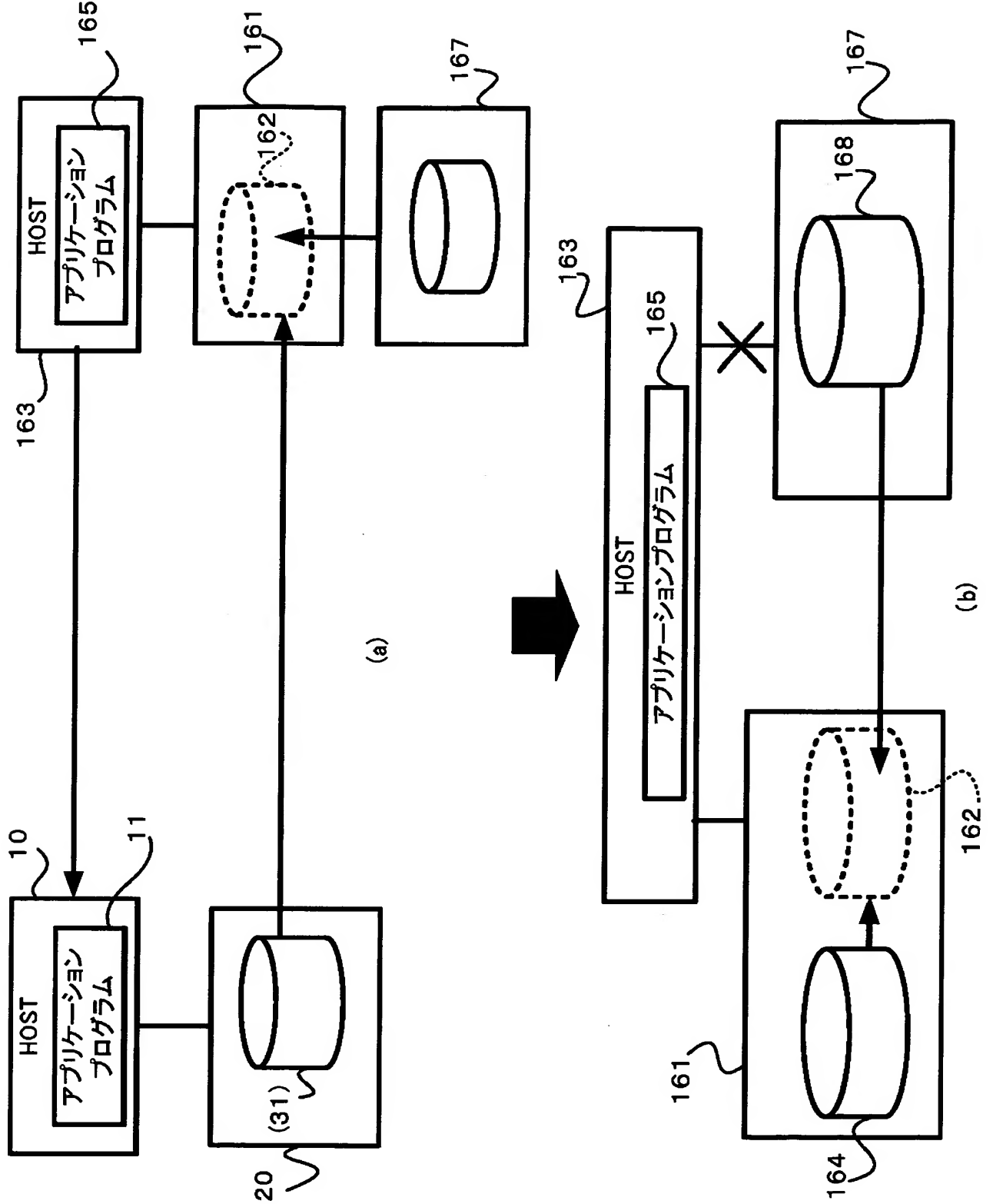
【図10】



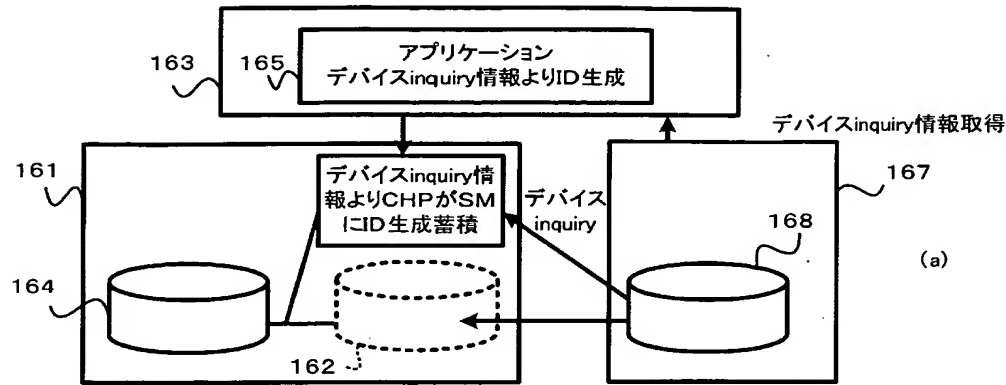
【図11】



【図12】



【図 13】



標準 inquiry データ抜粋

0	
4	
8	VendorID
12	
16	ProductID
20	
24	
28	
32	
36	Vendor Specific
40	
44	
48	
52	
56	
60	
64	
68	
72	
76	
80	

デバイス  
inquiry 情報から  
外部 LUN に対  
して unique な  
部分を抜粋し  
てデバイス識  
別情報 (ID) と  
してマッピング  
テーブルに格  
納する。  
例として DF で  
は VendorID,  
ProductID,  
VendorSpecific  
内のシリアル  
番号、LDEV  
番号を用いる

外部 LUN と内部 VDEV のマッピングテーブル

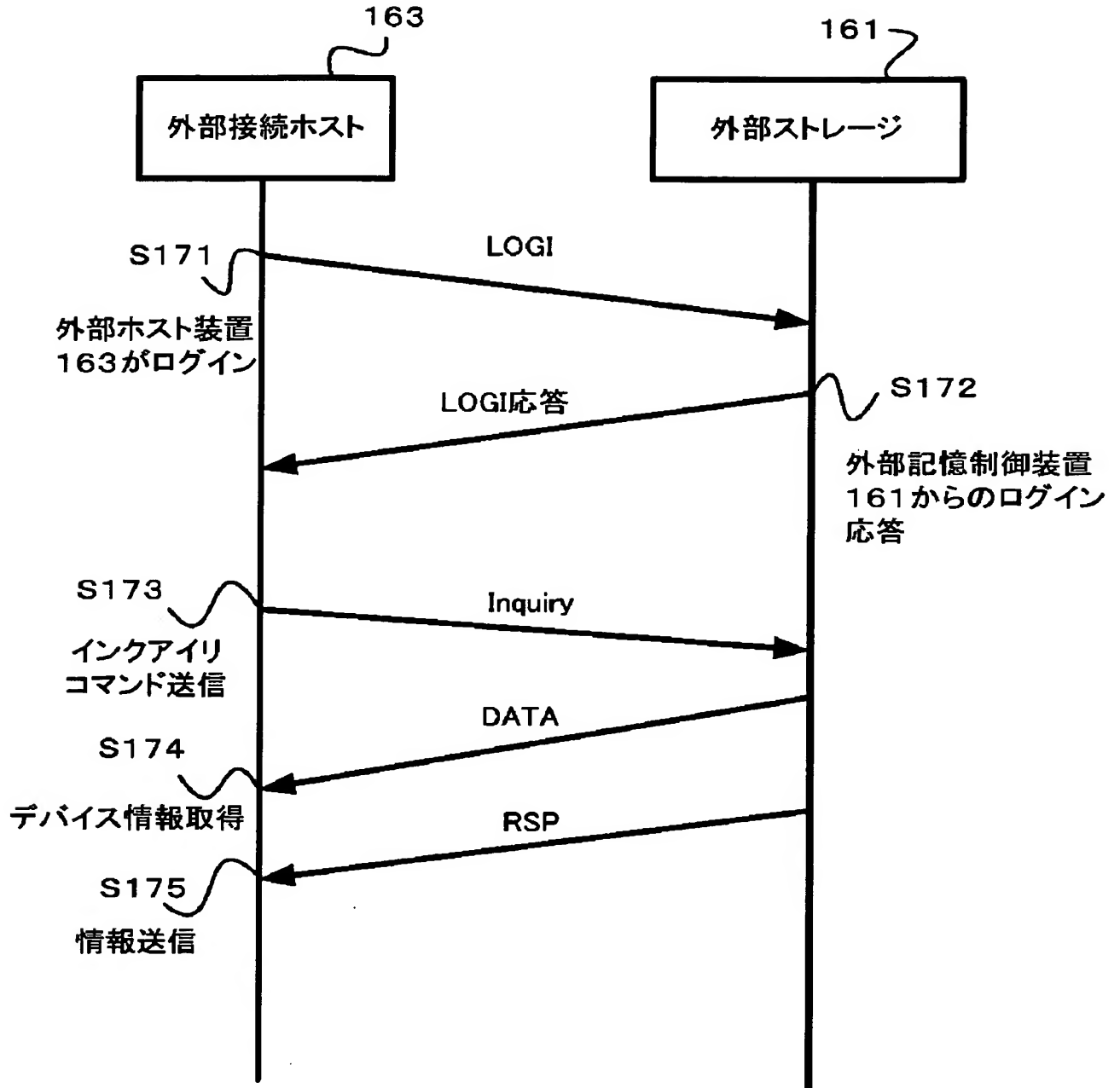
V D E V	外部デバイス情報				
	デバイス識 別情報 (ID)	Capacity (Kbyte)	デバイ ス種別	バス情報	
				WWN	LUX
0	DRFGTFNEIEK WNF1GKEF	657,456	DISK	0xAABBCCDD EEFF0000	0
1	ADRFGTNEIE KWN33445	89,854	DISK	0xAABBCCD DEEFF0000	3
2	GRRFFDDEF TT00	-	Tape	0x44556677 AA331122	5
3	AABBCCDD EEFF	5,544,223	Disk	0x77DE12435 5889922	6
				0x77DE12435 588910	3
				0x77DE12435 588988	5

(b)

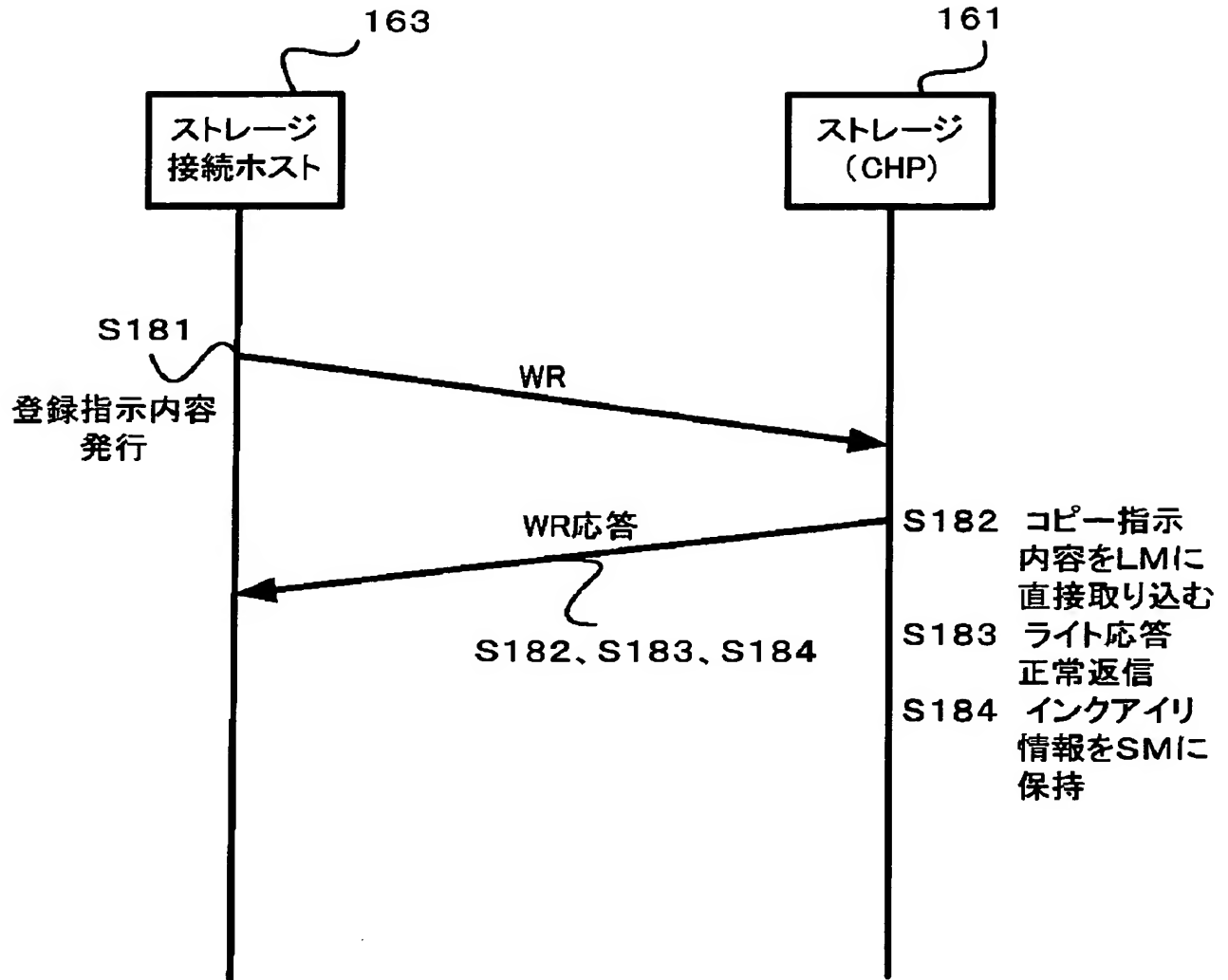
(c)



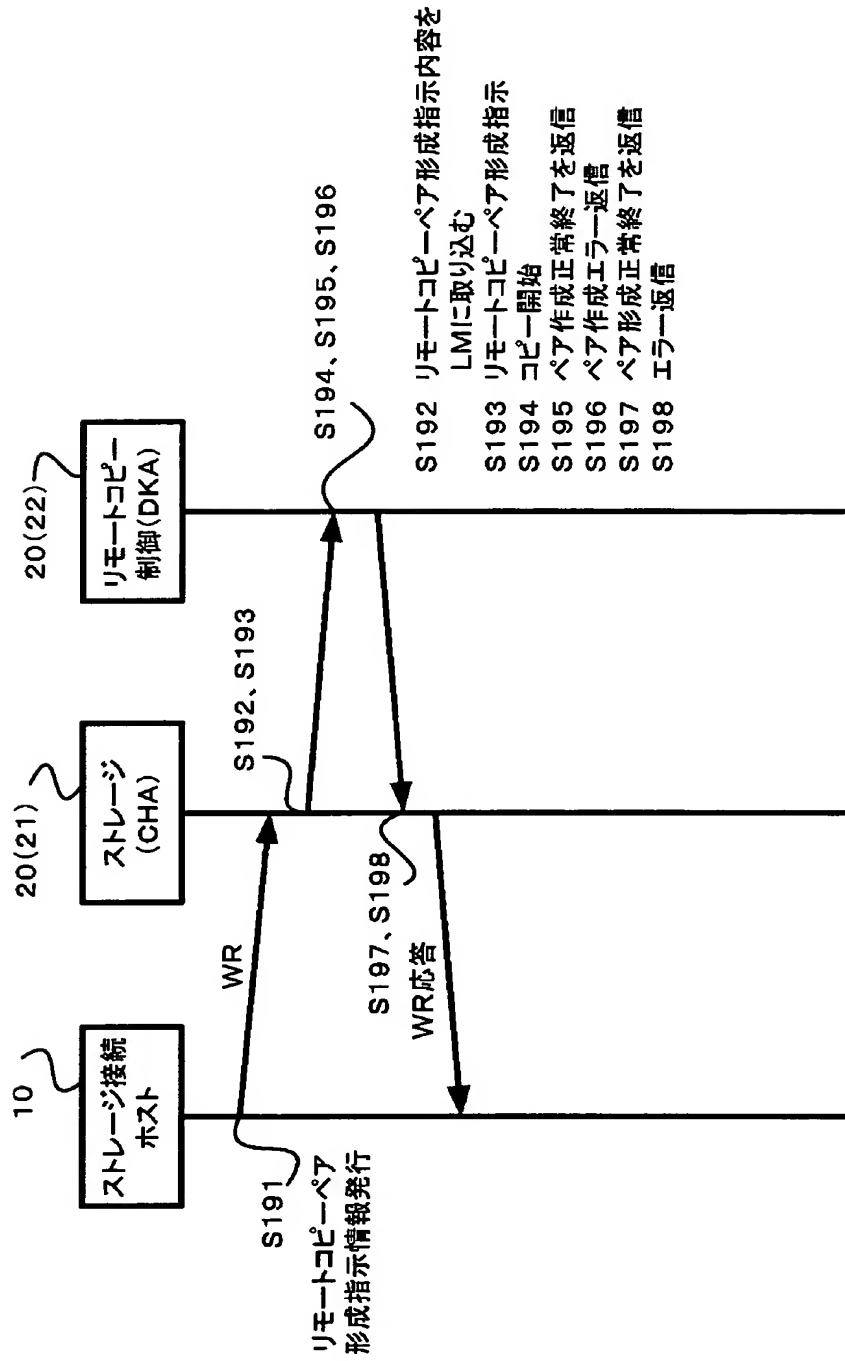
【図14】



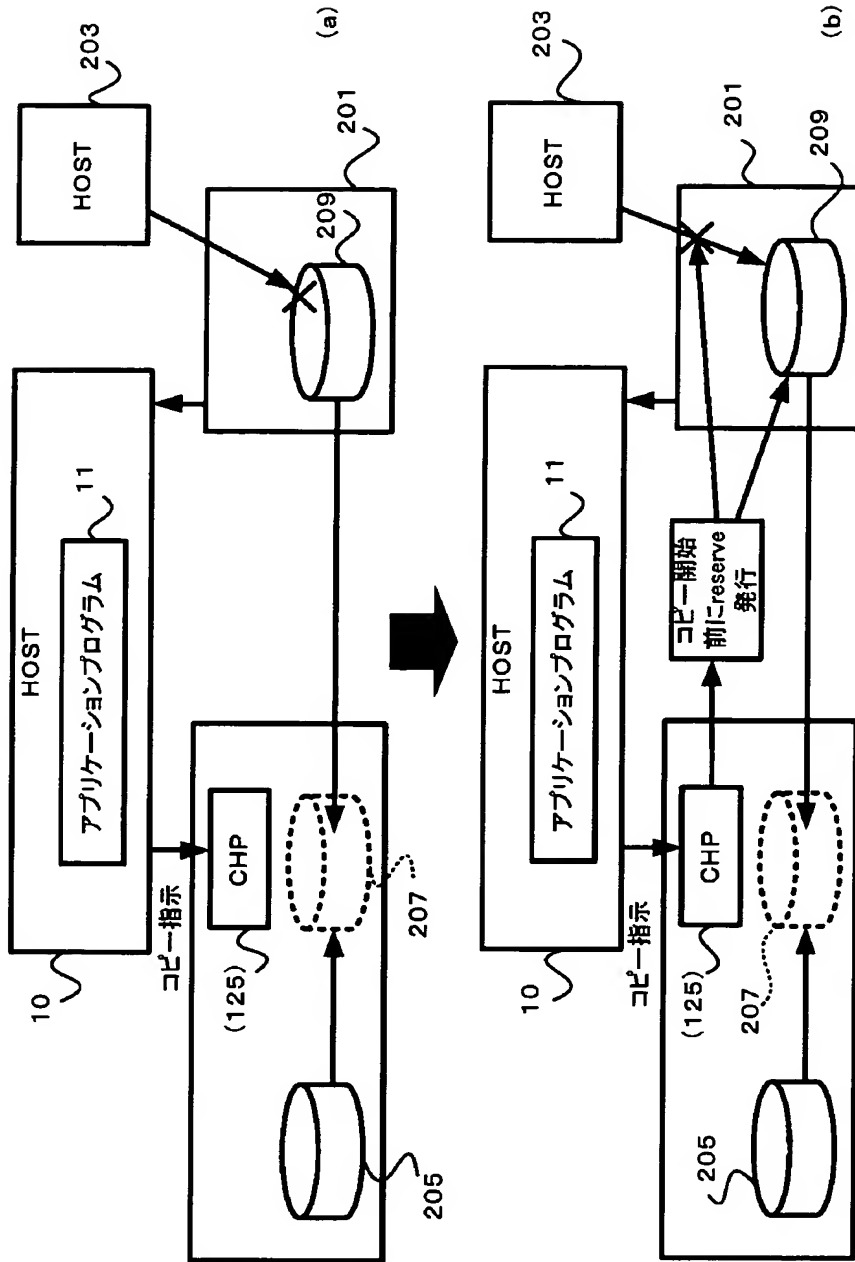
【図 15】



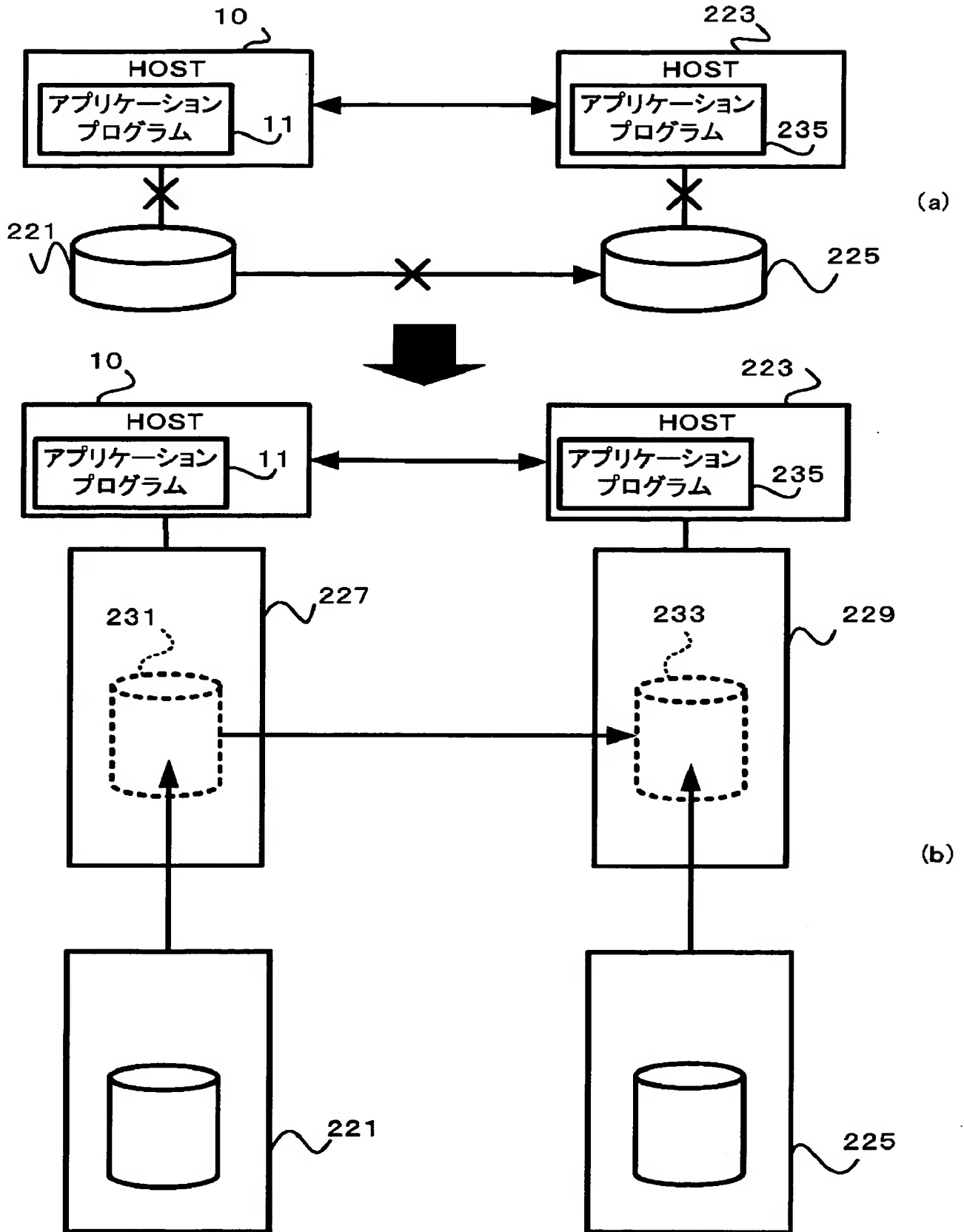
【図16】



【図 17】



【図18】



【書類名】 要約書

【要約】

【課題】 或るディスクアレイ装置から見て、外部に存在するディスクアレイ装置内の記憶デバイスを、上位装置と接続可能なデバイスの資源として利用できるようにする。

【解決手段】 チャンネルプロセッサ125は、アプリケーションプログラム11から、コピー指示内容が第1の記憶制御装置10へ転送されると、指示内容をローカルメモリ127に取り込む。指示内容の解析を行い解析結果に従ってオープンリモートコピー/MRCFの指示を、オープンリモートコピー/MRCF制御部129に出力する。指示内容に問題がなければ、ホスト装置10にライト応答を正常とする旨の返信を行う。ホスト装置10によるインクアイリ情報のリードに対処するため、共有メモリ25からインクアイリ情報を検索し、第1の記憶制御装置20からのホスト装置10への出力データを作成する。

【選択図】 図10

認定・付加情報

特許出願の番号	特願2003-400549
受付番号	50301971209
書類名	特許願
担当官	第七担当上席 0096
作成日	平成15年12月 1日

<認定情報・付加情報>

【提出日】 平成15年11月28日

特願 2003-400549

出願人履歴情報

識別番号

[000005108]

1. 変更年月日

1990年 8月31日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台4丁目6番地

氏 名

株式会社日立製作所